University of Padua

DEPARTMENT OF MATHEMATICS "Tullio Levi-Civita"

Master Degree in Mathematics

Existence and Uniqueness of Solutions for a class of McKean-Vlasov Stochastic Differential Equations

Supervisor:
Prof. Markus Fischer

Candidate:
Ofelia Bonesini
Student Number:
1179915

19 July 2019

Academic Year: 2018-2019

Contents

Introduction

In this thesis we deal with a class of McKean-Vlasov Stochastic Differential Equations (MV-SDEs).
MV-SDEs are more involved than classical SDEs as their coefficients depend on the law of the solution itself. They are sometimes referred to as mean-field SDEs and were first studied by McKean in *A class of Markov Process associated with non linear parabolic equations* (1966). These equations describe the limiting behaviour of individual particles having diffusive dynamics and which interact with each other in a "mean field" sense.
In this thesis, we are interested in showing existence and uniqueness of solutions for a class of MV-SDEs that arise in the study of the Large Deviation Principle for weakly interacting Itô diffusions. One way of proving limit theorems of this type is through the so-called weak convergence approach, which requires, besides tightness and identification of the limit, uniqueness of solutions for a controlled version of the limit model.

Let $T > 0$ be a time horizon and $d, d_1 \in \mathbb{N}$. Let $\mathcal{P}(\mathbb{R}^d)$ be the space of probability measures on $\mathbb{R}^d$ and let $\mathcal{P}_1(\mathbb{R}^d)$ be the space of probability measures on $\mathbb{R}^d$ with finite moment of first order.
Let $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t\in[0,T]})$ be a stochastic basis satisfying the usual hypotheses endowed with a d_1-dimensional $(\mathcal{F}_t)$-Wiener Process $(W_t)_{t\in[0,T]}$.
The aim of this work is to investigate the cases in which the SDE has the following form:

$$\begin{aligned} dX(t) =& b(t, X, \text{Law}(X(t)))dt + \sigma(t, X, \text{Law}(X(t)))u(t)dt \\ &+ \sigma(t, X, \text{Law}(X(t)))dW_t, \end{aligned} \tag{0.1}$$

with initial condition $X(0) = X_0$ fixed and $u \in \mathcal{U}$, where $\mathcal{U}$ is the space of all $(\mathcal{F}_t)$-progressively measurable functions $w : [0, T] \times \Omega \to \mathbb{R}^{d_1}$ such that:

$$\mathbb{E}\left[\int_0^T |w(s)|^2 ds\right] < \infty.$$

Equation (0.1) is called *controlled McKean-Vlasov stochastic differential equation.* It is sometimes referred to as (controlled) non-linear SDE. This denomination is due to the non-linearity of its infinitesimal generator.

The complications in such an expression are not only due to the non-linear nature of the equation but also to the presence of the *control* $u(\cdot)$. We think of u as a control, although we are not selecting u. It is simply a given stochastic process with finite integral moment of second order.

There are different ways to deal with both the problem of Existence and that of Uniqueness of solutions. The approach depends on the assumptions we make on the various coefficients appearing in the equation.
In particular, the case in which the coefficients are globally Lipschitz and the control is not present, or it is present but bounded, has already been discussed in several works and one can quite easily show that the solution in this case exists and is unique.
The presence of an unbounded control complicates quite a bit our problem, at least for the Uniqueness problem. In fact, the introduction of an unbounded control, u, makes the actual drift coefficient only locally Lipschitz. As opposed to ordinary stochastic differential equations, for MV-SDEs it is possible to find counterexamples to uniqueness when the coefficients are only locally Lipschitz, see [13].

In order to deal with the problem of Existence of solutions for the equations above, we introduce a functional Ψ that associates to an element $\theta = (\theta_t)_{t\in[0,T]} \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$ the flow of measures corresponding to the marginal laws of the solution of the SDE:

$$dX(t) = b(t, X, \theta(t))dt + \sigma(t, X, \theta(t))u(t)dt + \sigma(t, X, \theta(t))dW_t.$$

Then Ψ is well defined under quite general assumptions on the coefficients b and σ and with initial condition $X(0) = X_0$ fixed.
A fixed point for $\Psi(\cdot)$ is the flow of measures associated to a solution for the Equation (0.1).
We exploit *Brouwer-Schauder-Tychonoff Theorem* to show that $\Psi(\cdot)$ admits at least one fixed point. We need to restrict ourself to consider a convex compact subset of $\mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$ to be in the hypotheses of the theorem. The key tool to obtain compactness is a version of *Ascoli Arzelà Theorem.*
Furthermore we need to show that the functional $\Psi(\cdot)$ is continuous.
In the case in which the coefficients b and σ are globally Lipschitz the continuity of $\Psi(\cdot)$ can be showed directly. When the coefficients b and σ are simply uniformly continuous, we are still able to show the continuity of $\Psi(\cdot)$ but we need to exploit the *Martingale Problem.*

Secondly, we develop the Uniqueness topic.
We obtain some positive results, imposing stronger assumptions on coefficients or controls.

In particular, we are able to show uniqueness in the case of finite moments of exponential order of the control $u(\cdot)$, that is: there exists a positive $c > 0$ such that

$$\mathbb{E}\left[e^{c\int_0^T |u(t)|^2 dt}\right] = D < \infty.$$

We also show Uniqueness of solutions in the case of "delayed volatility coefficients", that is: there exists $\delta > 0$ such that for all $t \in [0, T]$, all $\phi, \psi \in \mathcal{C}([0, T], \mathbb{R}^d)$, all $\mu \in \mathcal{P}(\mathbb{R}^d)$:

$$\sigma(t, \phi, \mu) = \sigma(t, \psi, \mu), \quad \text{once} \quad \phi(s) = \psi(s) \quad \forall s \in [0, t - \delta].$$

Chapter 1

Existence of solutions

1.1 Relative Compactness

Let us consider a *Polish space**, E, and a complete metric $r(\cdot, \cdot)$ on it.
We define $C = \mathcal{C}([0,T], E)$, the space of continuous functions from the interval $[0,T]$ to E, and define the distance between two points x and y in C as :

$$\rho(x, y) := \sup_{t \in [0,T]} r(x(t), y(t)).$$

We introduce the concept of *modulus of continuity* :

Definition 1.1. *Let $x \in \mathcal{C}([0,T], E)$, we define the* ***modulus of continuity*** *of x as:*

$$w_x(\delta) = w(x, \delta) = \sup_{|s-t| \leq \delta} r(x(s), x(t)), \qquad 0 < \delta \leq T.$$

Remark 1.2. *The definition of modulus of continuity given above let us deduce a necessary and sufficient condition for an x to be uniformly continuous on $[0,T]$:*

$$\lim_{\delta \to 0} w_x(\delta) = 0. \tag{1.1}$$

Any $x \in C$ satisfies (1.1).

Remark 1.3. *We can notice, exploiting the property of the* sup *of a difference*

*A *Polish space* is a separable completely metrizable topological space; that is, a space homeomorphic to a complete metric space that has a countable dense subset.

and triangle inequality, that:

$$
\begin{aligned}
&|w_x(\delta) - w_y(\delta)| = |\sup_{|s-t|\le\delta} r(x(s), x(t)) - \sup_{|s-t|\le\delta} r(y(s), y(t))| \\
&\le \sup_{|s-t|\le\delta} |r(x(s), x(t)) - r(y(s), y(t))| \\
&\le \sup_{|s-t|\le\delta} |r(x(s), x(t)) - r(x(t), y(s)) + r(x(t), y(s)) - r(y(s), y(t))| \\
&\le \sup_{|s-t|\le\delta} \Big\{|r(x(s), x(t)) - r(x(t), y(s))| + |r(x(t), y(s)) - r(y(s), y(t))|\Big\} \\
&\le \sup_{|s-t|\le\delta} \Big\{r(x(s), y(s)) + r(x(t), y(t))\Big\} \le 2\rho(x, y),
\end{aligned}
$$

which means that $w(x,\delta)$ is continuous in x for a fixed δ.

We recall the definition of *relative compactness*.

Definition 1.4. *A set A is* ***relatively compact*** *if $\bar{A}$, the closure of A, is compact. This fact is equivalent to the condition that each sequence in A contains a convergent subsequence (which may not lie in A)*

Here we present a generalized version of *Ascoli-Arzelà Theorem*, Theorem A2.1 in [9], which completely characterizes relative compactness in $\mathcal{C}([0,T], E)$.

Theorem 1.5 (Ascoli -Arzelà)**.** *Fix two metric spaces $[0,T]$ and E, where the interval $[0,T]$ is, obviously, compact and E is complete, and let D be dense in $[0,T]$. Then, the set $A \subset \mathcal{C}([0,T], E)$ is relatively compact if and only if:*

$$\pi_t A \text{ is relatively compact in } E, \text{ for every } t \in D \tag{1.2}$$

and

$$\lim_{\delta\to 0} \sup_{x\in A} w_x(\delta) = 0. \tag{1.3}$$

In that case, even $\bigcup_{t\in[0,T]} \pi_t A$ is relatively compact in E.

Remark 1.6. *Let us notice that the functions in A are by definition* equicontinuous *at the point $t_0 \in [0,T]$ if, as $t \to t_0$, $\sup_{x\in A} r(x(t), x(t_0)) \to 0$; and the condition (1.3) defines* uniform *equicontinuity over $[0,T]$ of the functions in A.*

Proof: If $\bar{A}$ is compact, (1.2) follows easily.
Let us recall a result.

Lemma 1.7. *If $f_n \searrow 0$ for each x and if each f_n is everywhere upper semicontinuous the convergence is uniform on each compact set.*

Since $w(x, n^{-1})$ is continuous in x and non increasing in n, if $\bar{A}$ is compact exploiting the theorem above we have that $\lim_{\delta\to 0} w_x(\delta) = 0$ holds uniformly on A, that is (1.3).
Now suppose that (1.2) and (1.3) hold.
The idea now is to exploit (1.2) and (1.3) to prove that A is totally bounded. In fact, $C = \mathcal{C}([0,T], E)$ is complete, because of the completeness of E, and so A totally bounded is a condition equivalent to $\bar{A}$ compact.
Given $\epsilon > 0$, by hypothesis (1.3), we can choose $k \in \mathbb{N}$ large enough that $w_x(Tk^{-1}) < \epsilon$, for all $x \in A$.
$D \subset [0,T]$ is a dense subset of $[0,T]$. We can consider a finite set $I_k = \{t_1, \ldots, t_m\} \subset D$ such that $\forall t \in [0,T]$: $\exists t_j \in I_k$ such that $|t - t_j| \leq Tk^{-1}$.
By hypothesis (1.2), we know that, for any $t_j \in I_k \subset D$, $\pi_{t_j} A$ is relatively compact. Because of the completeness of E and the result quoted above, $\pi_{t_j} A$ is totally bounded, $\forall t_j \in I_k$. Hence, for all $t_j \in I_k$, we can define a finite ϵ-net, H_j, on $\pi_{t_j} A$.
Define H as $H = \bigotimes_{j=1}^{m} H_j$.
Consider $y \in H$.
Choose an element x_y in A such that $\sup_{j\in\{1,\ldots,m\}} r(x_y(t_j), y(t_j)) \leq \epsilon$, if it exists, and otherwise select arbitrarily an element in $\mathcal{C}([0,T], E)$ that satisfies the same condition and set it to be x_y. We build the set B as $B = \{x_y : y \in H\}$. B is finite, since $\#B = \#H = \prod_{j\in\{1,\ldots,m\}} \#H_j < \infty$. We want to show that B is a finite $\tilde{\epsilon}$-net for A.
Let z be a generic element in A.
$\forall t_j \in I_k$ there must be $y_j \in H_j$, such that $r(y_j, z(t_j)) \leq \epsilon$. Hence, there exists $y \in H$: $\sup_{j\in\{1,\ldots,m\}} r(y_j, z(t_j)) \leq \epsilon$.
Let $x_y \in B$ be the element in A such that $\sup_{j\in\{1,\ldots,m\}} r(y_j, x_y(t_j)) \leq \epsilon$.
We have, for all $t \in [0,T]$:

$$
\begin{aligned}
r(x_y(t), z(t)) &\leq r(x_y(t), x_y(t_i)) + r(x_y(t_i), y(t_i)) + r(y(t_i), z(t_i)) + r(z(t_i), z(t)) \\
&\leq w_{x_y}(Tk^{-1}) + \epsilon + \epsilon + w_z(Tk^{-1}) \leq 4\epsilon,
\end{aligned}
$$

where $t_i \in I_k$ is such that $|t - t_i| \leq Tk^{-1}$.
Hence we have that:

$$\rho(x_y, z) = \sup_{t\in[0,T]} r(x_y(t), z(t)) \leq 4\epsilon.$$

Thus B is a finite $4\epsilon-$net for A and, since ϵ is arbitrarily small, we can deduce that A is totally bounded as desired.

□

Theorem 1.8. *Suppose that* $0 = t_0 < t_1 < \cdots < t_v = T$ *and*

$$\min_{1<i<v} (t_i - t_{i-1}) \geq \delta. \tag{1.4}$$

Then, for arbitrary x,

$$w_x(\delta) \leq 3 \max_{1\leq i\leq v} \sup_{t_{i-1}\leq s\leq t_i} r(x(s), x(t_{i-1})). \tag{1.5}$$

Remark 1.9. *We underline that (1.4) does not require $t_i - t_{i-1} \geq \delta$ for the extremals $i = 1$ or $i = v$.*

Proof: Let m be the maximum in (1.5) that is
$m = \max_{1\leq i\leq v} \sup_{t_{i-1}\leq s\leq t_i} r(x(s), x(t_{i-1}))$.
Let us notice that if $|s - t| \leq \delta$, then s and t must lie in the same interval $[t_{i-1}, t_i]$ or at least in adjacent ones. In the first case, if for example s and t are in the same interval $[t_{l-1}, t_l]$, then:

$$r(x(t), x(s)) \leq r(x(t), x(t_{l-1})) + r(x(t_{l-1}), x(s)) \leq 2m.$$

In the second one, if s and t are in adjacent intervals, $[t_{i-1}, t_i]$ and $[t_i, t_{i+1}]$ respectively, then

$$r(x(s), x(t)) \leq r(x(s), x(t_{i-1})) + r(x(t_{i-1}), x(t_i)) + r(x(t_i), x(t)) \leq 3m.$$

This implies that

$$w_x(\delta) = \sup_{|s-t|\leq\delta} r(x(s), x(t)) \leq 3m,$$

that is (1.5).

□

1.2 Fixed-point Theorems

Here we give a few definitions and results that are necessary to the fixed-point theorems we would like to exploit. For this part we refer to chapter 17 in [1].

Definition 1.10. *A* ***correspondence*** ϕ *from a set* X *to a set* Y *assigns to each* x *in* X *a subset* $\phi(x)$ *of* Y*.*
We write $\phi : X \twoheadrightarrow Y$ *to distinguish a correspondence from a function from* X *to* Y*.*

We wish to denote restrictions on the values of a correspondence.

A correspondence $\phi : X \twoheadrightarrow Y$ is said to be **non-empty-valued** if, for each $x \in X$, $\phi(x)$ is a non-empty subset of Y.
If Y is a topological space, then we say that the correspondence $\phi : X \twoheadrightarrow Y$ is **closed-valued** or **has closed values** if $\phi(x)$ is a closed set for each $x \in X$ and analogously we say that ϕ is **convex-valued** if for each $x \in X$ $\phi(x)$ is a convex subset of Y.

Just as functions have inverses, so do correspondences.

Definition 1.11. *The* ***upper inverse*** ϕ^u *(also called the strong inverse) of a subset* A *of* Y *is defined by*

$$\phi^u(A) = \{x \in X : \phi(x) \subset A\}$$

Definition 1.12. *A correspondence,* $\phi : X \twoheadrightarrow Y$*, from a topological space to a topological vector space is* ***upper demicontinuous*** *if the upper inverse of every open half space in* Y *is open in* X*.*

Definition 1.13. *A correspondence,* $\phi : X \twoheadrightarrow Y$*, between topological spaces* ***has closed graph****, if its graph* $Gr(\phi) = \{(x, y) \in X \times Y ; y \in \phi(x)\}$ *is a closed subset of* $X \times Y$*.*

Definition 1.14. *Let* A *be a subset of a set* X*. A* ***fixed point*** *of a function* $f : A \to X$ *is a point* $x \in A$ *satisfying* $f(x) = x$*. A* ***fixed point*** *of a correspondence* $\phi : A \twoheadrightarrow X$ *is a point* $x \in A$ *satisfying* $x \in \phi(x)$*.*

Definition 1.15. *Let* A *be a subset of a vector space* X*. A correspondence* $\phi : A \twoheadrightarrow X$ *is* ***inward pointing*** *if, for each* $x \in A$*, there exists some* $y \in \phi(x)$ *and* $\lambda > 0$ *satisfying* $x + \lambda(y - x) \in A$*.*

We are now ready to state the following theorems.

Theorem 1.16 (Halpern-Bergman)**.** *Let* K *be a non empty compact convex subset of a locally convex Hausdorff space* X*, and let* $\phi : K \twoheadrightarrow K$ *be an inward pointing upper hemicontinuous correspondence with non-empty closed convex values. Then,* ϕ *has a fixed point.*

Theorem 1.17 (Kakutani-Fan-Glicksberg)**.** *Let K be a non-empty compact convex subset of a locally convex Hausdorff space, and let the correspondence $\phi : K \twoheadrightarrow K$ have closed graph and non-empty convex values. Then, the set of fixed points of ϕ is compact and non empty.*

The next fixed point theorem is immediate from the fact that continuous function define upper hemicontinuous correspondences, but is stated separately for historical reasons. This is the result we will use.

Corollario 1.18 (Brouwer-Schauder-Tychonoff)**.** *Let K be a non-empty compact convex subset of a locally convex Hausdorff space, and let $f : K \longrightarrow K$ be a continuous function. Then, the set of fixed points of f is compact and non-empty.*

1.3 Existence in Wasserstein-1 metric

The topic of existence of solutions can be developed considering different metrics on the space $\mathcal{P}_1(\mathbb{R}^d)$, which is the space of probability measures on $\mathbb{R}^d$ with finite moments of first order.
We focus our interest on the Wasserstein-1 metric.

Definition 1.19. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$, we define the distance in* ***Wasserstein-1 metric*** *between μ and ν as:*

$$W_1(\mu,\nu) := \inf_{\gamma\in\Gamma(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y| d\gamma(x,y),$$

where $\Gamma(\mu,\nu)$ denotes the set of all measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν on the first and the second factor respectively.

According to *Kantorovich-Rubistein Theorem*, for example Theorem 1.14 in [15], this metric can also be rewritten as:

$$W_1(\mu,\nu) = \sup\left\{ \int_{\mathbb{R}^d} f(x)\mu(dx) - \int_{\mathbb{R}^d} f(x)\nu(dx) : Lip(f) \leq 1 \right\},$$

where $Lip(f)$ denotes the Lipschitz constant of f.
In particular, this definition justifies the following:

Proposition 1.20. *Setting μ and $\nu \in \mathcal{P}_1(\mathbb{R}^d)$, let X_μ (resp. X_ν) be a stochastic variable whose law is μ (resp. ν). It is true that:*

$$W_1(\mu,\nu) \leq \mathbb{E}\Big[|X_\mu - X_\nu|\Big].$$

Proof. Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a Lipschitz function such that its Lipschitz constant $\mathrm{Lip}(f) \leq 1$. This implies that $\forall x, y \in \mathbb{R}^d$, it holds:

$$|f(x) - f(y)| \leq \mathrm{Lip}(f)|x-y| \leq |x-y|.$$

Hence, we have:

$$\int_{\mathbb{R}^d} f(x)\mu(dx) - \int_{\mathbb{R}^d} f(y)\nu(dy) = \mathbb{E}[f(X_\mu)] - \mathbb{E}[f(X_\nu)] \leq |\mathbb{E}[f(X_\mu) - f(X_\nu)]|$$
$$\leq \mathbb{E}\big[|f(X_\mu) - f(X_\nu)|\big] \leq \mathbb{E}\big[|X_\mu - X_\nu|\big].$$

This means that:

$$W_1(\mu,\nu) = \sup\left\{ \int_{\mathbb{R}^d} f(x)\mu(dx) - \int_{\mathbb{R}^d} f(x)\nu(dx) : Lip(f) \leq 1 \right\} \leq \mathbb{E}\big[|X_\mu - X_\nu|\big].$$

□

Remark 1.21. *Let us highlight that we have the natural inclusion $\mathcal{P}_1(\mathbb{R}^d) \subset \mathcal{X}$, where $\mathcal{X}$ denotes the space of finite signed measures endowed with the topology of weak convergence.*†
Furthermore, we can notice that, if $(\eta_n)_{n\in\mathbb{N}}$ is a convergent sequence in $(\mathcal{P}_1(\mathbb{R}^d)$,$W_1)$, it converges also in $(\mathcal{X}, d_{bL})$. The converse is not true. Let $(\gamma_n)_{n\in\mathbb{N}} \subset \mathcal{P}_1(\mathbb{R}^d)$ be a convergent sequence in $(\mathcal{X}, d_{bL})$. In general, this sequence does not converge in $(\mathcal{P}_1(\mathbb{R}^d), W_1)$. In fact, convergence in $(\mathcal{P}_1(\mathbb{R}^d), W_1)$ is stronger than the one in $(\mathcal{X}, d_{bL})$, because it requires the convergence of the first moments.

We introduce the space of continuous functions on $\mathcal{P}_1(\mathbb{R}^d)$, $\mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$, equipped with the following metric, induced by Wasserstein-1:

Definition 1.22. *Let $\mu, \nu \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$, we define the metric $\rho_1(\cdot,\cdot)$ as:*

$$\rho_1(\mu,\nu) := \sup_{t\in[0,T]} W_1(\mu_t,\nu_t)$$

Remark 1.23. *Let $\mathcal{C}([0,T], \mathcal{X})$ be the space of continuous functions on the space of finite signed measures, $\mathcal{X}$, equipped with the following metric:*

$$\rho_{bL}(\mu,\nu) := \sup_{t\in[0,T]} d_{bL}(\mu_t,\nu_t), \qquad \mu,\nu \in \mathcal{X}.$$

$\mathcal{C}([0,T], \mathcal{X})$ is a locally convex Hausdorff topological vector space.
There is a natural inclusion: $\mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d)) \subset \mathcal{C}([0,T], \mathcal{X})$.
Furthermore, the topologies, induced by the metrics ρ_1, ρ_{bL}, are compatible.
Let $(f_n)_{n\in\mathbb{N}}$ be a convergent sequence in $(\mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d)), \rho_1)$. This sequence does converge in $(\mathcal{C}([0,T], \mathcal{X}), \rho_{bL})$. The converse is not true.
This implies that any compact subset of $\mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$ is a compact subset of $\mathcal{C}([0,T], \mathcal{X})$. These facts will be helpful in the future.

Now, we have to recall the background of our problem.

Let $T > 0$ be the time horizon and $d, d_1 \in \mathbb{N}$.
Let $\mathcal{U}$ be the set of all quadruples $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t\in[0,T]}, u, W)$ such that the pair $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t\in[0,T]})$ forms a stochastic basis satisfying the usual hypotheses, W is a d_1-dimensional $(\mathcal{F}_t)$-Wiener process and u is an $\mathbb{R}^{d_1}$-valued, $(\mathcal{F}_t)$-progressively measurable process such that:

$$\mathbb{E}\Big[\int_0^T |u(t)|^2 dt\Big] = Q < \infty.$$

For simplicity, we may write $u \in \mathcal{U}$ instead of $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t\in[0,T]}, u, W) \in \mathcal{U}$.
Let b, σ be predictable functionals on $[0,T] \times \mathcal{C}([0,T], \mathbb{R}^d) \times \mathcal{P}_1(\mathbb{R}^d)$, with values

†Let us recall that the metric d_{bL} induces on $\mathcal{X}$ the topology of weak convergence. As a reference on can see Theorem (11.3.3) in [6].

in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d_1}$, respectively.
Given $u \in \mathcal{U}$, we would like to study the following *non-linear controlled SDE*:

$$\begin{aligned} dX_t =& b(t, X, \mathrm{Law}(X_t))dt + \sigma(t, X, \mathrm{Law}(X_t))u(t)dt \\ &+ \sigma(t, X, \mathrm{Law}(X_t))dW_t, \end{aligned} \tag{1.6}$$

where the initial condition is $X(0) = X_0$, with $\mathbb{E}[|X_0|^2] < \infty$.
A solution of (1.6) under $u \in \mathcal{U}$ is a continuous $\mathbb{R}^d$-valued process X defined on the given stochastic basis and adapted to the given filtration such that the integral version of (1.6) holds with probability one. A solution is said to be strong when it is adapted to the filtration $(\bar{\mathcal{F}}_t^{B,X_0})$, i.e. the filtration generated by W and X_0, completed with respect to $\mathbb{P}$. We say that for Equation (1.6) there is trajectorial uniqueness (we simply write uniqueness) if, given a couple of solutions X and X', defined on the same filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ with the same Wiener process W, the processes X, X' are indistinguishable, that is $\mathbb{P}(X(t) = X'(t), \forall t \in [0, T]) = 1$.

Consider the following Lipschitz and growth conditions on b and σ.

(L) There exists L such that for all $t \in [0, T]$, all $\phi, \psi \in \mathcal{C}([0, T], \mathbb{R}^d)$ all $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$

$$|b(t, \phi, \mu) - b(t, \psi, \nu)| + |\sigma(t, \phi, \mu) - \sigma(t, \psi, \nu)| \le L\Big(\sup_{s\in[0,t]} |\phi(s) - \psi(s)| + W_1(\mu, \nu)\Big).$$

(G) There exists a constant $K > 0$ such that for all $t \in [0, T]$, all $\phi \in \mathcal{C}([0, T], \mathbb{R}^d)$, all $\mu \in \mathcal{P}_1(\mathbb{R}^d)$

$$|b(t, \phi, \mu)| \le K\Big(1 + \sup_{s\in[0,t]} |\phi(s)|\Big), \qquad |\sigma(t, \phi, \mu)| \le K.$$

Remark 1.24. *We want to underline that these conditions are sufficient to guarantee that, for any fixed $\theta \in \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$, the following SDE has a unique strong solution, whose moment of second order is bounded by a constant that does not depend on the specific θ:*[‡]

$$\begin{aligned} &dX_t = b(t, X, \theta_t)dt + \sigma(t, X, \theta_t)u(t)dt + \sigma(t, X, \theta_t)dW_t \\ &X(0) = X_0. \end{aligned} \tag{1.7}$$

As we have anticipated in the Introduction, we will deduce Existence of solutions for the Equation (1.6) with a fixed point argument.
To this purpose, we introduce the following functional:

$$\Psi : \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d)) \longrightarrow \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$$

[‡]See the Appendix B for the proof, in particular Theorem B.1.

that associates to $\nu \in \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$ $\Psi(\nu) = (\Psi(\nu)_t)_{t\in[0,T]} \in \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$, such that $\Psi(\nu)_t = \text{Law}(X^\nu(t)), \forall t \in [0,T]$, where X^ν is the unique solution of the SDE (1.7), with $\theta = \nu$.
Thanks to Remark 1.24, we already know that this functional is well defined.
A fixed point for the application Ψ is a $\mu \in \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$ such that $\Psi(\mu) = \mu$, that is equivalent to the fact that $\mu(t) = \text{Law}(X^\mu(t)), \forall t \in [0,T]$, where X^μ is the solution of the SDE:

$$\begin{aligned} &dX_t = b(t,X,\mu_t)dt + \sigma(t,X,\mu_t)u(t)dt + \sigma(t,X,\mu_t)dW_t \\ &X(0) = X_0, \end{aligned} \tag{1.8}$$

hence μ is the law of a solution of McKean-Vlasov SDE.
In particular, μ is the law of the solution of Equation (1.8), namely X^μ.
By hypothesis, Equation (1.7), for a fixed flux of measures μ, has a unique strong solution. Hence, we are able to deduce the existence of a strong solution for McKean-Vlasov SDE.

Remark 1.25. *The argument that we will develop is not sufficient to have the uniqueness of solution. In fact, the fixed-point theorem that we will exploit guarantees the existence of at least a fixed point, not its uniqueness.*

To show that the function Ψ admits at least a fixed-point, we would like to exploit *Brouwer-Schauder-Tychonoff Theorem*.
We need to prove that the hypotheses of the theorem are satisfied.

First of all, we need to focus our attention on a non-empty, compact convex subset of the space $\mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$, in order to restrict $\Psi(\cdot)$ to this set. Let us notice that, according to the Remark 1.23, we only need to find such a compact subset in $\mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$; it will then naturally be a compact subset of the locally convex Hausdorff space $\mathcal{C}([0,T],\mathcal{X})$.

Define $A \subset \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$ as:

$$\begin{aligned} A := \{\mu \in \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d)) : \mu(t) = \text{Law}(X(t)), \forall t \in [0,T], \\ \text{where } X = (X(t))_{t\in[0,T]} \text{ is the solution to the following equation:} \\ dX_t = b(t,X,\theta_t)dt + \sigma(t,X,\theta_t)u(t)dt + \sigma(t,X,\theta_t)dW_t, \\ \text{with initial condition } X(0) = X_0, \text{ fixed, and } \theta \in \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))\}. \end{aligned}$$

Obviously, A is non-empty and, thanks to Remark 1.24, we have already noticed that each $\mu \in A$ is the law of a process $(X_t)_{t\in[0,T]}$, such that:

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^2\Big] \le Re^{8K^2T^2} = R',$$

with $R = 4(\mathbb{E}[|X_0|^2] + TK^2\mathbb{E}[\int_0^T |u(t)|^2 dt] + 4TK^2 + 2K^2T^2)$.

Now, consider $B \subset \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$.

$$B := \{\mu \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d)) : \mu(t) = \text{Law}(Y(t)), \quad t \in [0,T], \text{ with}$$
$$Y(t) = X_0 + \int_0^t \hat{b}_s ds + \int_0^t \hat{\sigma}_s dW_s,$$
where $\hat{b}_s$ is a $\mathbb{R}^d$-valued, $\mathcal{F}_t$-progressively measurable process, such that:
$$\mathbb{E}[\int_s^t |\hat{b}_r| dr] \leq C\sqrt{t-s}, \qquad 0 \leq s \leq t \leq T,$$
$\hat{\sigma}_s$ is a $\mathbb{R}^{d\times d_1}$-valued, $\mathcal{F}_t$-progressively measurable process, bounded by the constant K and
$$\mathbb{E}\Big[\sup_{t\in[0,T]} |Y(t)|^2\Big] \leq R'\},$$

where $C = K((2T + 2TR')^{\frac{1}{2}} + \mathbb{E}[\int_0^T |u(t)|^2 dt]^{\frac{1}{2}}) < \infty$.

Furthermore, we can prove the following.

Proposition 1.26. *Consider the subsets of* $\mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$*, A and B defined above. Then the following inclusion holds:*

$$A \subset B.$$

Proof. We have to show that any $\mu \in A$ is in B.
Consider a $\mu \in A$. It must be $\mu(t) = \text{Law}(X^\theta(t)), \forall t \in [0,T]$, where X^θ is the solution of Equation (1.7), for a fixed $\theta \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$.
We must check that drift and volatility coefficients, in this case, satisfy the hypotheses to let μ be in B. Certainly, b, σ and u take values in the right space and are $\mathcal{F}_t$-progressively measurable, by hypothesis.
By Remark 1.24 we already know that

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X^\theta|^2\Big] \leq R'.$$

Furthermore, by hypothesis (G) on σ, we know that: $|\sigma(t,\phi,\mu)| \leq K$, for any $\phi \in \mathcal{C}([0,T], \mathbb{R}^d)$, $\mu \in \mathcal{P}_1(\mathbb{R}^d)$ and $t \in [0,T]$.
To conclude that $A \subset B$, we have only to show the hypothesis on the drift term.

Exploiting triangle inequality and Hölder inequality, we can deduce:

$$\begin{aligned}
&\mathbb{E}\Big[\int_s^t |b(r,X^\theta,\theta(r))+\sigma(r,X^\theta,\theta(r))u(r)|dr\Big]\\
&\leq \mathbb{E}\Big[\int_s^t |b(r,X^\theta,\theta(r))|+|\sigma(r,X^\theta,\theta(r))u(r)|dr\Big]\\
&\leq \mathbb{E}\Big[\int_s^t |b(r,X^\theta,\theta(r))|dr\Big]+\mathbb{E}\Big[\int_s^t |\sigma(r,X^\theta,\theta(r))u(r)|dr\Big]\\
&\leq \mathbb{E}\Big[|\int_s^t |b(r,X^\theta,\theta(r))|dr|^2\Big]^{\frac{1}{2}}+\mathbb{E}\Big[|\int_s^t |\sigma(r,X^\theta,\theta(r))u(r)|dr|^2\Big]^{\frac{1}{2}}=\triangle.
\end{aligned}$$

Exploiting (G) hypotheses on b and σ and Hölder inequality, we have:

$$\begin{aligned}
&\triangle \leq \mathbb{E}\Big[|\int_s^t K(1+\sup_{w\in[0,r]}|X^\theta(w)|)dr|^2\Big]^{\frac{1}{2}}+\mathbb{E}\Big[|\int_s^t K|u(r)|dr|^2\Big]^{\frac{1}{2}}\\
&\leq K\mathbb{E}\Big[(t-s)\int_s^t (1+\sup_{w\in[0,r]}|X^\theta(w)|)^2 dr\Big]^{\frac{1}{2}}+K\mathbb{E}\Big[(t-s)\int_s^t |u(r)|^2 dr\Big]^{\frac{1}{2}}\\
&\leq \sqrt{2}K\sqrt{t-s}\,\mathbb{E}\Big[\int_s^t 1+\sup_{w\in[0,r]}|X^\theta(w)|^2 dr\Big]^{\frac{1}{2}}+K\sqrt{t-s}\,\mathbb{E}\Big[\int_0^T |u(r)|^2 dr\Big]^{\frac{1}{2}}\\
&\leq K\sqrt{t-s}\Bigg(\Bigg(2T+2\int_0^T \mathbb{E}\Bigg[\sup_{w\in[0,r]}|X^\theta(w)|^2\Bigg]dr\Bigg)^{\frac{1}{2}}+\mathbb{E}\Big[\int_0^T |u(r)|^2 dr\Big]^{\frac{1}{2}}\Bigg)\\
&\leq K\sqrt{t-s}\Bigg(\Big(2T+2TR'\Big)^{\frac{1}{2}}+\mathbb{E}\Big[\int_0^T |u(r)|^2 dr\Big]^{\frac{1}{2}}\Bigg)=C\sqrt{t-s}.
\end{aligned}$$

Hence, we can conclude that $A \subset B$. □

Remark 1.27. *In particular, this implies the following chain of inclusions* $\mathrm{Im}(\Psi) = A \subset B \subset \bar{B}$.
Hence, the following restriction of Ψ *is well defined:*

$$\Psi : \bar{B} \longrightarrow \bar{B}.$$

Furthermore, we can notice that B *is non-empty. In fact,* A *is non-empty and* $A \subset B$.

We will prove that $\bar{B}$, the closure of B, is the non-empty, compact, convex subset of $\mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$, which we have to restrict Ψ to, in order to be in the hypothesis of *Brouwer-Schauder-Tychonoff Theorem.*

Proposition 1.28. *Consider the non-empty set* B*, which is defined as:*

$B := \{\mu \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d)) : \mu(t) = Law(Y(t)), \quad t \in [0,T],$ *with*

$Y(t) = X_0 + \int_0^t \hat{b}_s ds + \int_0^t \hat{\sigma}_s dW_s,$

where $\hat{b}_s$ *is a* $\mathbb{R}^d$*-valued,* $\mathcal{F}_t$*-progressively measurable process, such that:*

$\mathbb{E}[\int_s^t |\hat{b}_r| dr] \leq C\sqrt{t-s}, \qquad 0 \leq s \leq t \leq T,$

$\hat{\sigma}_s$ *is a* $\mathbb{R}^{d \times d_1}$*-valued,* $\mathcal{F}_t$*-progressively measurable process, bounded by the constant* K *and*

$\mathbb{E}\Big[\sup_{t\in[0,T]} |Y(t)|^2\Big] \leq R'\},$

where $C = K((2T + 2TR')^{\frac{1}{2}} + \mathbb{E}[\int_0^T |u(t)|^2 dt]^{\frac{1}{2}}) < \infty$.
Then, $\bar{B}$*, the closure of* B*, is compact in* $\mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$.

Proof. The statement of the proposition is equivalent to the relative compactness of B.
In order to show that B is relatively compact, we would like to exploit a general version of *Ascoli-Arzelà Theorem*, namely Theorem 1.5.
First, we would like to show that there exists a set D dense in $[0,T]$ such that $\forall t \in D$: $\pi_t(B)$ is relatively compact in $\mathcal{P}_1(\mathbb{R}^d)$. Actually, in this case we will have $D = [0,T]$. By definition of B, we already know that there exists a positive constant, $R' < \infty$, such that for any process Y such that $\mu_Y \in B : \mathbb{E}[\sup_{t\in[0,T]} |Y(t)|^2] \leq R'$.
Let us notice that for any process Y such that $\mu_Y \in B$, exploiting Markov inequality, we have, $\forall t \in [0,T]$:

$$\mathbb{P}(|Y(t)| > \epsilon) = \mathbb{P}(|Y(t)|^2 > \epsilon^2) \leq \frac{\mathbb{E}[|Y(t)|^2]}{\epsilon^2} \leq \frac{R'}{\epsilon^2}$$

Hence, we can deduce that, for any fixed $t \in [0,T]$:

$$\sup_{\mu_Y(t)\in\pi_t(B)} \mathbb{P}(|Y(t)| \in \bar{B}_\epsilon(0)) \geq 1 - \frac{R'}{\epsilon^2}.$$

So, $\forall t \in [0,T]$, the set $\pi_t(B)$ is tight.
By Prokhorov Theorem, we have that $\pi_t(B)$ is relatively compact in $\mathcal{P}(\mathbb{R}^d)$.
Now, we would like to show that $\forall t \in [0,T]$ $\pi_t(B)$ is relatively compact in $\mathcal{P}_1(\mathbb{R}^d)$. In order to do this, we will exploit the notion of sequential compactness. Let us consider a sequence $(\mu_n)_{n\in\mathbb{N}} \subset cl(\pi_t(B))$. Since $cl(\pi_t(B))$ is compact in $\mathcal{P}(\mathbb{R}^d)$, there must be a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ of $(\mu_n)_{n\in\mathbb{N}}$ such that:

$$\mu_{n_k} \longrightarrow_{k\to\infty} \mu, \text{ in } \mathcal{P}(\mathbb{R}^d). \tag{1.9}$$

We would like to show that

$$\mu_{n_k} \longrightarrow_{k\to\infty} \mu, \text{ in } \mathcal{P}_1(\mathbb{R}^d), \tag{1.10}$$

in order to deduce that $cl(\pi_t(B))$ is compact in $(\mathcal{P}_1(\mathbb{R}^d), W_1)$, and so that $\pi_t(B)$ is relatively compact in $(\mathcal{P}_1(\mathbb{R}^d), W_1)$.[§]
By *Dominated Convergence Theorem*, for instance Lemma 3.11 in [9], we have that (1.9) implies (1.10), once the sequence $(X_{n_k})_{k\in\mathbb{N}}$, with $\mu_{n_k} = \text{Law}(X_{n_k})$, is uniformly integrable.
Uniform integrability is a condition equivalent to $\sup_{k\in\mathbb{N}} \mathbb{E}[|X_{n_k}| \cdot \mathbb{I}_{\{|X_{n_k}|>M\}}] \to 0$, as $M \to \infty$.
$\forall k \in \mathbb{N}$, exploiting Hölder inequality and finiteness of moments of second order of the elements of the sequence, we have:

$$\mathbb{E}[|X_{n_k}| \cdot \mathbb{I}_{\{|X_{n_k}|>M\}}] \le \sqrt{\mathbb{E}[|X_{n_k}|^2]} \cdot \sqrt{\mathbb{E}[\mathbb{I}_{\{|X_{n_k}|>M\}}]} \le \sqrt{R'}\sqrt{\mathbb{P}(|X_{n_k}| > M)}$$
$$\le \sqrt{R'}\sqrt{\mathbb{P}(|X_{n_k}|^2 > M^2)} \le \sqrt{R'}\sqrt{\frac{\mathbb{E}[|X_{n_k}|^2]}{M^2}} \le \sqrt{R'}\sqrt{\frac{R'}{M^2}} = \frac{R'}{M},$$

that clearly goes to zero as M goes to infinity.
We proved the first hypothesis of Ascoli-Arzelà Theorem.

In order to deduce our thesis, we have to prove the following:

$$\lim_{\delta\to 0} \sup_{\mu\in B} w_\mu(\delta) = 0,$$

where $w_\mu(\cdot)$ denotes the modulus of continuity of μ, that, in this case, is defined as:

$$w_\mu(\delta) = \sup_{|s-t|\le\delta} W_1(\mu(s), \mu(t)), \quad \text{with } \delta \in (0, T].$$

Now, fix $\mu \in B$ and $\delta \in (0, T]$. Let $0 \le s < t \le T$ such that $|s - t| \le \delta$. Suppose that μ is the law of a certain process $(Y(t))_{t\in[0,T]}$, whose coefficients satisfy the hypotheses to let μ be in B.

Exploiting Proposition 1.20, the hypotheses on Y and Hölder inequality, we have that:

$$W_1(\mu(s), \mu(t)) \le \mathbb{E}\Big[|Y(s) - Y(t)|\Big] = \mathbb{E}\Big[|\int_s^t \hat{b}_r dr + \int_s^t \hat{\sigma}_r dW_r|\Big]$$
$$\le \mathbb{E}\Big[|\int_s^t \hat{b}_r dr|\Big] + \mathbb{E}\Big[|\int_s^t \hat{\sigma}_r dW_r|\Big] \le \mathbb{E}\Big[\int_s^t |\hat{b}_r| dr\Big] + \mathbb{E}\Big[|\int_s^t \hat{\sigma}_r dW_r|^2\Big]^{\frac{1}{2}} = \triangle.$$

[§]Here we exploited the fact that Wasserstein-1 metric corresponds to weak convergence plus convergence of the first moments topology, see Theorem 7.12 in [15].

Exploiting the hypotheses on $\hat{b}$ and $\hat{\sigma}$ and Itô Isometry, we can deduce:

$$\triangle \leq C\sqrt{t-s} + \mathbb{E}\Big[\int_s^t |\hat{\sigma}_r|^2 dr\Big]^{\frac{1}{2}} \leq C\sqrt{t-s} + \mathbb{E}\Big[\int_s^t K^2 dr\Big]^{\frac{1}{2}} \leq (C+K)\sqrt{t-s}.$$

This estimation does not depend on the specific t and s, but only on their difference, namely $\delta = |t-s|$. Furthermore, it is valid for each $\mu \in B$.

Hence, we can deduce the following estimation:

$$\sup_{\mu \in B} w_\mu(\delta) = \sup_{\mu \in B} \sup_{|t-s| \leq \delta} W_1(\mu(t), \mu(s)) \leq \sup_{\mu \in B} \sup_{|t-s| \leq \delta} (C+K)\sqrt{t-s}$$
$$\leq \sup_{\mu \in B} (C+K)\sqrt{\delta} = (C+K)\sqrt{\delta}.$$

This shows that:

$$0 \leq \lim_{\delta \to 0} \sup_{\mu \in B} w_\mu(\delta) \leq \lim_{\delta \to 0} (C+K)\sqrt{\delta} = 0,$$

and this ends our proof. □

In the second place, we need to show that the closure of B is convex.
In general, the convexity of B implies the convexity of its closure. We can limit ourself to show that B is convex.

Proposition 1.29. *Consider the non-empty set B, which is defined as:*

$B := \{\mu \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d)) : \mu(t) = Law(Y(t)), \quad t \in [0,T],$ *with*

$$Y(t) = X_0 + \int_0^t \hat{b}_s ds + \int_0^t \hat{\sigma}_s dW_s,$$

where $\hat{b}_s$ is a $\mathbb{R}^d$-valued, $\mathcal{F}_t$-progressively measurable process, such that:

$$\mathbb{E}[\int_s^t |\hat{b}_r| dr] \leq C\sqrt{t-s}, \qquad 0 \leq s \leq t \leq T$$

$\hat{\sigma}_s$ is a $\mathbb{R}^{d \times d_1}$-valued, $\mathcal{F}_t$-progressively measurable process, bounded by the constant K and

$$\mathbb{E}\Big[\sup_{t \in [0,T]} |Y(t)|^2\Big] \leq R'\},$$

where $C = K((2T + 2TR')^{\frac{1}{2}} + \mathbb{E}[\int_0^T |u(t)|^2 dt]^{\frac{1}{2}}) \leq \infty$.
Then, B is convex.

Proof. Let μ and ν be a couple of elements in B. Suppose that, $\forall t \in [0,T]$, $\mu(t)$ =Law$(Y(t))$ and $\nu(t)$ =Law$(Z(t))$, respectively, with:

$$Y(t) = X_0 + \int_0^t \hat{b}_s ds + \int_0^t \hat{\sigma}_s dW_s,$$
$$Z(t) = X_0 + \int_0^t \breve{b}_s ds + \int_0^t \breve{\sigma}_s dW_s,$$

with suitable hypotheses on coefficients.
Our purpose is to show that any convex combination of the type $\eta_\lambda = \lambda\mu + (1-\lambda)\nu$, for $\lambda \in [0,1]$ is an element of B.
Let ξ_λ be a Bernoullian variable of parameter λ, independent from the processes in B and $\mathcal{F}_0$-measurable.[¶] For any $t \in [0,T]$, $\eta_{\lambda t}$ is the law of the process $W_\lambda(t) = \xi_\lambda \cdot Y(t) + (1-\xi_\lambda)\cdot Z(t)$. In fact, for any $D \in \mathcal{F}$, for any $t \in [0,T]$, we have:

$$\begin{aligned}
\eta_{\lambda t}(D) &= \mathbb{P}(W_\lambda(t) \in D) = \mathbb{P}(\xi_\lambda \cdot Y(t) + (1-\xi_\lambda)\cdot Z(t) \in D)\\
&= \mathbb{P}(\xi_\lambda \cdot Y(t) + (1-\xi_\lambda)\cdot Z(t) \in D | \xi_\lambda = 1)\cdot \mathbb{P}(\xi_\lambda = 1)\\
&\quad + \mathbb{P}(\xi_\lambda \cdot Y(t) + (1-\xi_\lambda)\cdot Z(t) \in D | \xi_\lambda = 0)\cdot \mathbb{P}(\xi_\lambda = 0)\\
&= \mathbb{P}(Y(t) \in D)\cdot\lambda + \mathbb{P}(Z(t)\in D)\cdot(1-\lambda) = \lambda\mu_t(D) + (1-\lambda)\nu_t(D).
\end{aligned}$$

We have to check the hypotheses on drift and volatility coefficients of W_λ, in order to conclude that $\eta_\lambda \in B$. We can rewrite $W_\lambda(t)$, for $t \in [0,T]$ as:

$$\begin{aligned}
&W_\lambda(t) = \xi_\lambda \cdot Y(t) + (1-\xi_\lambda)\cdot Z(t)\\
&= \xi_\lambda \cdot \Big(X_0 + \int_0^t \hat{b}_s ds + \int_0^t \hat{\sigma}_s dW_s\Big) + (1-\xi_\lambda)\cdot\Big(X_0 + \int_0^t \breve{b}_s ds + \int_0^t \breve{\sigma}_s dW_s\Big)\\
&= X_0 + \int_0^t \xi_\lambda \cdot \hat{b}_s + (1-\xi_\lambda)\cdot \breve{b}_s ds + \int_0^t \xi_\lambda\cdot\hat{\sigma}_s + (1-\xi_\lambda)\cdot\breve{\sigma}_s dW_s\\
&= X_0 + \int_0^t \dot{b}_s ds + \int_0^t \dot{\sigma}_s dW_s
\end{aligned}$$

Naturally, $\dot{b}_s$ and $\dot{\sigma}_s$ take values in the right spaces and are both $\mathcal{F}_t$-progressively measurable. Furthermore, $\forall s \in [0,T]$, we have that

$$\begin{aligned}
|\dot{\sigma}_s| &= |\xi_\lambda\cdot\hat{\sigma}_s + (1-\xi_\lambda)\cdot\breve{\sigma}_s| \le |\xi_\lambda|\cdot|\hat{\sigma}_s| + |(1-\xi_\lambda)|\cdot|\breve{\sigma}_s|\\
&\le |\xi_\lambda|\cdot K + |(1-\xi_\lambda)|\cdot K = \xi_\lambda\cdot K + (1-\xi_\lambda)\cdot K = K
\end{aligned}$$

and that

$$\begin{aligned}
&\mathbb{E}\Big[\int_s^t |\dot{b}_r| dr\Big] = \mathbb{E}\Big[\int_s^t |\xi_\lambda\cdot\hat{b}_r + (1-\xi_\lambda)\cdot\breve{b}_r| dr\Big]\\
&= \mathbb{E}\Big[\int_s^t |\xi_\lambda\cdot\hat{b}_r + (1-\xi_\lambda)\cdot\breve{b}_r| dr\Big|\xi_\lambda = 1\Big]\cdot\mathbb{P}(\xi_\lambda = 1)\\
&\qquad + \mathbb{E}\Big[\int_s^t |\xi_\lambda\cdot\hat{b}_r + (1-\xi_\lambda)\cdot\breve{b}_r| dr\Big|\xi_\lambda = 0\Big]\cdot\mathbb{P}(\xi_\lambda = 0)\\
&= \mathbb{E}\Big[\int_s^t |\hat{b}_r| dr\Big]\cdot\lambda + \mathbb{E}\Big[\int_s^t |\breve{b}_r| dr\Big]\cdot(1-\lambda) \le C\sqrt{t-s}\cdot\lambda + C\sqrt{t-s}\cdot(1-\lambda)\\
&= C\sqrt{t-s}.
\end{aligned}$$

[¶] If it is not possible to define such a variable ξ_λ, we can extend the sample space Ω in order to make it feasible.

Finally, we have:

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|W(t)|^2\Big] = \mathbb{E}\Big[\sup_{t\in[0,T]}|\xi_\lambda \cdot Y(t) + (1-\xi_\lambda)\cdot Z(t)|^2\Big]$$
$$= \mathbb{E}\Big[\sup_{t\in[0,T]}|Y(t)|^2\Big]\mathbb{P}(\xi_\lambda = 1) + \mathbb{E}\Big[\sup_{t\in[0,T]}|Z(t)|^2\Big]\mathbb{P}(\xi_\lambda = 0)$$
$$\leq \lambda R' + (1-\lambda)R' = R'.$$

This ends the proof that $\eta_\lambda \in B$. Hence, B is convex. □

There is one last step to end our proof of the existence of solutions for Equation (1.6).
Once we prove the continuity of the functional Ψ, we are in the hypotheses of Brouwer-Schauder-Tychonoff Theorem and we get the existence of solutions.

Proposition 1.30. *The functional Ψ, introduced previously, is continuous with respect to the metric on $\mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$, defined by:*
$\rho_1(\mu,\nu) = \sup_{t\in[0,T]} W_1(\mu_t,\nu_t)$.

Proof: Let μ be a fixed element in $\mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$ and in the same space consider a sequence $(\mu_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty}\rho_1(\mu,\mu_n) = 0$.
We have to prove that this implies that $\lim_{n\to\infty}\rho_1(\Psi(\mu),\Psi(\mu_n)) = 0$, in order to deduce the continuity of Ψ.
Exploiting the definition of $\rho_1(\cdot,\cdot)$ and Proposition 1.20, we have that:

$$\rho_1(\Psi(\mu),\Psi(\mu_n)) = \sup_{t\in[0,T]} W_1(\Psi(\mu)_t,\Psi(\mu_n)_t) \leq \sup_{t\in[0,T]}\mathbb{E}[|X_t^\mu - X_t^{\mu_n}|]$$
$$\leq \mathbb{E}[\sup_{t\in[0,T]}|X_t^\mu - X_t^{\mu_n}|] = \star,$$

where $(X_t^\mu)_{t\in[0,T]}$ (resp.$(X_t^{\mu_n})_{t\in[0,T]}$) denotes the solution to the SDE
$dX_t = b(t,X,\mu_t)dt + \sigma(t,X,\mu_t))u(t)dt + \sigma(t,X,\mu_t)dW_t,\ X(0) = X_0$
(resp. $dX_t = b(t,X,\mu_n(t))dt + \sigma(t,X,\mu_n(t))u(t)dt + \sigma(t,X,\mu_n(t))dW_t$).
We would like to show that the right term ($\star$) goes to zero, as n goes to infinity, to infer our thesis.

For $M \in \mathbb{N}$, define an $(\mathcal{F}_t)$-stopping time τ_M by

$$\tau_M(\omega) = \inf\{t\in[0,T] : \int_0^t |u(s,\omega)|^2 ds \geq M\}$$

with $\inf\varnothing = \infty$. Observe that $\mathbb{P}(\tau_M \leq T) \longrightarrow 0$, as $M\to\infty$,
since $\mathbb{E}[\int_0^T |u(t)|^2 dt] < \infty$.

Using Hölder's inequality, Doob's maximal inequality, Itô Isometry we obtain for $M \in \mathbb{N}$ and all $t \in [0,T]$:

$$\mathbb{E}\Bigg[\sup_{s\in[0,t]} |X^{\mu}_{s\wedge\tau_M} - X^{\mu_n}_{s\wedge\tau_M}|^2\Bigg] \leq 4T\mathbb{E}\Bigg[\int_0^{t\wedge\tau_M} |b(r,X^{\mu},\mu(r)) - b(r,X^{\mu_n},\mu_n(r))|^2 dr\Bigg]$$

$$+ 4\mathbb{E}\Bigg[\int_0^{t\wedge\tau_M} |\sigma(r,X^{\mu},\mu(r)) - \sigma(r,X^{\mu_n},\mu_n(r))|^2 dr. \int_0^{t\wedge\tau_M} |u(r)|^2 dr\Bigg]$$

$$+ 16\mathbb{E}\Bigg[\int_0^{t\wedge\tau_M} |\sigma(r,X^{\mu},\mu(r)) - \sigma(r,X^{\mu_n},\mu_n(r))|^2 dr\Bigg]$$

$$\leq 4T\mathbb{E}\Bigg[\int_0^{t\wedge\tau_M} |b(r,X^{\mu},\mu(r)) - b(r,X^{\mu_n},\mu_n(r))|^2 dr\Bigg]$$

$$+ (4M+16)\mathbb{E}\Bigg[\int_0^{t\wedge\tau_M} |\sigma(r,X^{\mu},\mu(r)) - \sigma(r,X^{\mu_n},\mu_n(r))|^2 dr\Bigg] = \triangle.$$

Now, we exploit hypothesis (L) to deduce the following:

$$\triangle \leq 8L^2(T+M+4)\mathbb{E}\Bigg[\int_0^{t\wedge\tau_M} \sup_{s\in[0,r]} |X^{\mu}_s - X^{\mu_n}_s|^2 + W_1(\mu(r),\mu_n(r))^2 dr\Bigg]$$

$$\leq 8L^2(T+M+4)\Big(\mathbb{E}\Big[\int_0^{t\wedge\tau_M} \sup_{s\in[0,r]} |X^{\mu}_s - X^{\mu_n}_s|^2 dr\Big] + \sup_{r\in[0,T]} W_1(\mu(r),\mu_n(r))^2\Big)$$

$$\leq 8L^2(T+M+4)\int_0^t \mathbb{E}\Big[\sup_{s\in[0,r]} |X^{\mu}_{s\wedge\tau_M} - X^{\mu_n}_{s\wedge\tau_M}|^2\Big] dr + 8L^2(T+M+4)T\rho_1(\mu,\mu_n).$$

Applying Gronwall Lemma we have:

$$\mathbb{E}\Bigg[\sup_{s\in[0,t]} |X^{\mu}_{s\wedge\tau_M} - X^{\mu_n}_{s\wedge\tau_M}|^2\Bigg] \leq 8L^2(T+M+4)T\rho_1(\mu,\mu_n)e^{8L^2(T+M+4)t},$$

and we can deduce that $\forall M \in \mathbb{N}$:

$$\mathbb{E}\Bigg[\sup_{s\in[0,T]} |X^{\mu}_{s\wedge\tau_M} - X^{\mu_n}_{s\wedge\tau_M}|^2\Bigg] \leq 8L^2(T+M+4)T\rho_1(\mu,\mu_n)e^{8L^2(T+M+4)T}.$$

Now, we go back to our first case.
Exploiting Cauchy-Schwarz inequality, linearity of expected values, Hölder inequality and the fact that:
$\sup\Bigg\{\mathbb{E}\Big[\sup_{t\in[0,T]} |X^{\mu}_s|^2\Big], \sup_{n\in\mathbb{N}} \mathbb{E}\Big[\sup_{t\in[0,T]} |X^{\mu_n}_s|^2\Big]\Bigg\} = R' < \infty$, for any

$M \in \mathbb{N}$ we have that:

$$\begin{aligned}
\mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_s^\mu - X_s^{\mu_n}|\Bigg] &= \mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_s^\mu - X_s^{\mu_n}|(\mathbb{I}_{\tau_M > T} + \mathbb{I}_{\tau_M \le T})\Bigg] \\
&= \mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_s^\mu - X_s^{\mu_n}| \cdot \mathbb{I}_{\tau_M > T}\Bigg] + \mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_s^\mu - X_s^{\mu_n}| \cdot \mathbb{I}_{\tau_M \le T}\Bigg] \\
&\le \mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_{s\wedge\tau_M}^\mu - X_{s\wedge\tau_M}^{\mu_n}|\Bigg] + \mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_s^\mu - X_s^{\mu_n}| \cdot \mathbb{I}_{\tau_M \le T}\Bigg] \\
&\le \mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_{s\wedge\tau_M}^\mu - X_{s\wedge\tau_M}^{\mu_n}|^2\Bigg]^{\frac{1}{2}} + \mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_s^\mu - X_s^{\mu_n}|^2\Bigg]^{\frac{1}{2}} \mathbb{P}\Big(\tau_M \le T\Big)^{\frac{1}{2}} \\
&\le 2L\sqrt{2(T+M+4)T\rho_1(\mu,\mu_n)}e^{4L^2(T+M+4)T} + 2\sqrt{R'}\mathbb{P}\Big(\tau_M \le T\Big)^{\frac{1}{2}}
\end{aligned}$$

Since $\lim_{M\to\infty}\mathbb{P}(\tau_M \le T) = 0$, we can choose $\bar{M} \in \mathbb{N}$ such that $\mathbb{P}(\tau_{\bar{M}} < T) \le \frac{\epsilon^2}{16R'}$.
Placing $\bar{M}$ in the inequality above, we can deduce:

$$\mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_s^\mu - X_s^{\mu_n}|\Bigg] \le 2L\sqrt{2(T+\bar{M}+4)T\rho_1(\mu,\mu_n)}e^{4L^2(T+\bar{M}+4)T} + \frac{\epsilon}{2}.$$

Hence, for any $\epsilon > 0$, we can find an $\tilde{n} \in \mathbb{N}$ such that $\forall n \ge \tilde{n}$:

$$\rho_1(\mu,\mu_n) \le \frac{1}{32L^2(T+\bar{M}+4)T}e^{-8L^2(T+\bar{M}+4)T}\epsilon^2,$$

and we have that:

$$\rho_1(\Psi(\mu),\Psi(\mu_n)) \le \mathbb{E}\Bigg[\sup_{s\in[0,T]}|X_s^\mu - X_s^{\mu_n}|\Bigg] \le \frac{\epsilon}{2} + \frac{\epsilon}{2} \le \epsilon.$$

□

Finally, we would like to show that for a sequence $(\mu_n)_{n\in\mathbb{N}} \subset B$, $\mu_n \to_{n\to\infty} \mu$ in $\mathcal{C}([0,T],\mathcal{P}(\mathbb{R}^d))$ (w.r.t. ρ_{bL}) implies $\mu_n \to_{n\to\infty} \mu$ in $\mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$ (w.r.t. ρ_1).
This fact allows us to deduce that the continuity of the functional $\Psi|_B$ w.r.t. ρ_1 implies its continuity w.r.t. ρ_{bL}.

We have the following conditions:

i)

$$\rho_{bL}(\mu_n,\mu) = \sup_{t\in[0,T]} d_{bL}(\mu_n(t),\mu(t)) \longrightarrow_{n\to\infty} 0,$$

which implies that

$$\sup_{t\in[0,T]}\Big|\int M\wedge|x|\mu_n(t,dx)-\int M\wedge|x|\mu(t,dx)\Big|\longrightarrow_{n\to\infty}0,\forall M\in\mathbb{N}.$$

ii)

$$\sup\Big\{\sup_{n\in\mathbb{N}}\sup_{t\in[0,T]}\int|x|^2\mu_n(t,dx),\ \sup_{t\in[0,T]}\int|x|^2\mu(t,dx)\Big\}\leq R'<\infty.$$

To deduce our thesis we need to show the uniform convergence of the first moments, that is:

$$\sup_{t\in[0,T]}\Big|\int|x|\mu_n(t,dx)-\int|x|\mu(t,dx)\Big|\longrightarrow_{n\to\infty}0.$$

We can write:

$$\sup_{t\in[0,T]}\Big|\int|x|\mu_n(t,dx)-\int|x|\mu(t,dx)\Big|=\sup_{t\in[0,T]}\Big|\int|x|(\mu_n(t,dx)-\mu(t,dx))\Big|$$
$$=\sup_{t\in[0,T]}\Big|\int_{|x|>M}|x|(\mu_n(t,dx)-\mu(t,dx))\Big|+\sup_{t\in[0,T]}\Big|\int_{|x|\leq M}|x|(\mu_n(t,dx)-\mu(t,dx))\Big|$$
$$=A+B.$$

Exploiting Hölder inequality, Markov inequality and condition ii), we can write:

$$A\leq\sup_{t\in[0,T]}\int_{|x|>M}|x|\mu_n(t,dx)+\sup_{t\in[0,T]}\int_{|x|>M}|x|\mu(t,dx)$$
$$\leq\sup_{t\in[0,T]}\int_{|x|>M}\mu_n(t,dx)^{\frac{1}{2}}\int|x|^2\mu_n(t,dx)^{\frac{1}{2}}+\sup_{t\in[0,T]}\int_{|x|>M}\mu(t,dx)^{\frac{1}{2}}\int|x|^2\mu(t,dx)^{\frac{1}{2}}$$
$$\leq\sqrt{R'}\Bigg(\sup_{t\in[0,T]}\int_{|x|>M}\mu_n(t,dx)^{\frac{1}{2}}+\sup_{t\in[0,T]}\int_{|x|>M}\mu(t,dx)^{\frac{1}{2}}\Bigg)\leq\frac{2R'}{M},$$

that could be made smaller than $\frac{\epsilon}{3}$ choosing $\bar{M}\geq\frac{6R'}{\epsilon}$.

Analogously, we have that

$$
\begin{aligned}
B = &\sup_{t\in[0,T]} \Big| \int_{|x|\le M} |x| \wedge M(\mu_n(t,dx) - \mu(t,dx)) \Big| \\
\le &\sup_{t\in[0,T]} \Big| \int |x| \wedge M(\mu_n(t,dx) - \mu_(t,dx)) \Big| \\
&\quad + \sup_{t\in[0,T]} \Big| \int_{|x|>M} |x| \wedge M(\mu_n(t,dx) - \mu(t,dx)) \Big| \\
\le &\sup_{t\in[0,T]} \Big| \int |x| \wedge M(\mu_n(t,dx) - \mu(t,dx)) \Big| \\
&\quad + M \sup_{t\in[0,T]} \Big| \int_{|x|>M} (\mu_n(t,dx) - \mu(t,dx)) \Big| \\
\le &\sup_{t\in[0,T]} \Big| \int |x| \wedge M(\mu_n(t,dx) - \mu(t,dx)) \Big| \\
&\quad + M \sup_{t\in[0,T]} \Big(\int_{|x|>M} \mu_n(t,dx) + \int_{|x|>M} \mu(t,dx)) \Big) \\
\le &\sup_{t\in[0,T]} \Big| \int |x| \wedge M(\mu_n(t,dx) - \mu(t,dx)) \Big| + 2M\frac{R'}{M^2}.
\end{aligned}
$$

Now, we can make $+2M\frac{R'}{M^2} = \frac{2R'}{M}$ smaller than $\frac{\epsilon}{3}$ by a proper choice of $\bar{M}$, as above. Furthermore, by hypothesis i) for a fixed $\bar{M}$ we can choose $\bar{n}$ such that: $\forall n \ge \bar{n}$ $\sup_{t\in[0,T]} \Big| \int |x| \wedge \bar{M}(\mu_n(t,dx) - \mu(t,dx)) \Big| \le \frac{\epsilon}{3}$. Hence, we have that $\forall n \ge \bar{n}$:

$$
\sup_{t\in[0,T]} \Big| \int |x|\mu_n(t,dx) - \int |x|\mu(t,dx) \Big| \le \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon,
$$

and, since $\epsilon > 0$ is arbitrary, we can conclude.

1.4 Existence under relaxed hypotheses

Looking at the proof of Existence of solutions for Equation (1.6) in previous section, we can notice that we used hypothesis (L) only in the proof of the continuity of the functional $\Psi(\cdot)$ and to guarantee that Equation (1.7) has a unique strong solution for all fixed $\theta \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$.
In this section we will show that hypothesis (L) can be relaxed and that we are able to guarantee Existence of solutions for Equation (1.6) under the hypotheses that $b(\cdot)$ and $\sigma(\cdot)$ are uniformly continuous and that the SDE (1.7) has a unique strong solution for $\theta \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$, with finite moments of second order.

As above, let $T > 0$ be the time horizon and $d, d_1 \in \mathbb{N}$.
Let $\mathcal{U}$ be the set of all quadruples $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t\in[0,T]}, u, W)$ such that the pair $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t\in[0,T]})$ forms a stochastic basis satisfying the usual hypothesis, W is a d_1-dimensional $(\mathcal{F}_t)$-Wiener process and u is an $\mathbb{R}^{d_1}$-valued, $(\mathcal{F}_t)$-progressively measurable process such that:

$$\mathbb{E}\Big[\int_0^T |u(t)|^2 dt\Big] = Q < \infty.$$

For simplicity, we may write $u \in \mathcal{U}$ instead of $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t\in[0,T]}, u, W) \in \mathcal{U}$.
Let b, σ be predictable functionals on $[0,T] \times \mathcal{C}([0,T], \mathbb{R}^d) \times \mathcal{P}_1(\mathbb{R}^d)$, with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d_1}$, respectively.
Given $u \in \mathcal{U}$, we consider the *non-linear controlled SDE*:

$$\begin{aligned} dX_t =& b(t, X, \mathrm{Law}(X_t))dt + \sigma(t, X, \mathrm{Law}(X_t))u(t)dt \\ &+ \sigma(t, X, \mathrm{Law}(X_t))dW_t, \end{aligned} \tag{1.11}$$

where the initial condition is $X(0) = X_0$, with $\mathbb{E}\big[|X_0|^2\big] < \infty$.
In this section we establish the following conditions on the coefficients b and σ and on the control u:

(UC) The functions $b(t,\cdot,\cdot)$ and $\sigma(t,\cdot,\cdot)$ are continuous, uniformly in $t \in [0,T]$.

(G) There exists a constant $K > 0$ such that for all $t \in [0,T]$, all $\phi \in \mathcal{C}([0,T], \mathbb{R}^d)$, all $\mu \in \mathcal{P}_1(\mathbb{R}^d)$

$$|b(t,\phi,\mu)| \le K\Big(1 + \sup_{s\in[0,t]} |\phi(s)|\Big), \qquad |\sigma(t,\phi,\mu)| \le K.$$

(U) The coefficients b, σ and the control u are such that for all $\theta \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$ the Equation below admits a unique strong solution with moments of second order bounded by the constant $R' > 0$, i.e. $\mathbb{E}[\sup_{t\in[0,T]} |X_t|^2] \le R'$.

$$\begin{aligned} &dX_t = b(t, X, \theta_t)dt + \sigma(t, X, \theta_t)u(t)dt + \sigma(t, X, \theta_t)dW_t \\ &X(0) = X_0. \end{aligned} \tag{1.12}$$

Remark 1.31. *We want to underline that the conditions (UC) and (G) are not sufficient to guarantee existence and uniqueness of solution for the equation above for any fixed* $\theta \in \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$*. Hence, to exploit the same results used in the previous section it is necessary to add hypothesis (U).*

As anticipated above, the only part that needs to be changed because of this new set of hypotheses is the proof of the continuity of the functional $\Psi(\cdot)$.
Let us recall that the functional, $\Psi : \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d)) \to \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$, associates to $\theta \in \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$ $\Psi(\theta) \in \mathcal{C}([0,T],\mathcal{P}_1(\mathbb{R}^d))$, such that $\Psi(\theta)_t = \mathrm{Law}(X^\theta(t)), \forall t \in [0,T]$, where X^θ is the unique strong solution of the SDE (1.12).
Thanks to the hypothesis (U), we already know that this functional is well defined.
We will exploit the *Martingale Problem* to deduce that $\Psi(\cdot)$ is a continuous functional.

First, we need to enlarge the space where our solutions live.
It will be advantageous to have a path space which is Polish for the control process $u \in \mathcal{U}$. We decide to work with the space of deterministic relaxed controls on $\mathbb{R}^{d_1} \times [0,T]$ with finite first moments. Let us first recall some facts about deterministic relaxed controls (see, e.g.,[3]). Let $\mathcal{R}$ denote the space of all deterministic relaxed controls on $\mathbb{R}^{d_1} \times [0,T]$, that is, $\mathcal{R}$ is the set of all positive measures r on $\mathcal{B}(\mathbb{R}^{d_1} \times [0,T])$ such that $r(\mathbb{R}^{d_1} \times [0,t]) = t$ for all $t \in [0,T]$. Let $r \in \mathcal{R}$ and $B \in \mathcal{B}(\mathbb{R}^{d_1})$. Then, the mapping $[0,T] \ni t \mapsto r(B \times [0,t])$ is absolutely continuous, hence differentiable almost everywhere.
$\mathcal{B}(\mathbb{R}^{d_1})$ is countably generated. Hence, the time derivative of r exists almost everywhere and is a measurable mapping $r_t : [0,T] \to \mathcal{P}(\mathbb{R}^{d_1})$ such that $r(dy \times dt) = r_t(dy)dt$. Let $\mathcal{R}_1$ denote the space of deterministic relaxed controls with finite first moments, that is,

$$\mathcal{R}_1 := \left\{ r \in \mathcal{R} : \int_{\mathbb{R}^{d_1}\times[0,T]} |y| r(dy \times dt) < \infty \right\}.$$

By definition, $\mathcal{R}_1 \subset \mathcal{R}$. $\mathcal{R}$ endowed with the topology of weak convergence of measures is a Polish space (not compact in our case). We equip $\mathcal{R}_1$ with the topology of weak convergence of measures plus convergence of first moments. This topology turns $\mathcal{R}_1$ into a Polish space (cf.[3]). We can notice that, for $T = 1$ (else one has to renormalize), the topology coincides with that induced by the Kantorovich–Rubinstein distance or Wasserstein distance of order one. Since the controls appear in an unbounded (but affine) fashion in the dynamics, ordinary weak convergence will not imply convergence of corresponding integrals, but convergence in $\mathcal{R}_1$ will.
Any $\mathbb{R}^{d_1}$-valued process v defined on some probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ induces

an $\mathcal{R}$-valued random variable ς according to

$$\varsigma_\omega(B \times I) := \int_I \delta_{v(t,\omega)}(B)dt, \quad B \in \mathcal{B}(\mathbb{R}^{d_1}), I \subset [0,T], \omega \in \Omega'.$$

If v is such that $\int_0^T |v(t,\omega)|dt < \infty$ for all $\omega \in \Omega'$, then the induced random variable ς takes values in $\mathcal{R}_1$. If v is progressively measurable with respect to a filtration $(\mathcal{F}'_t)$ in $\mathcal{F}$, then ς is adapted in the sense that the mapping $t \mapsto \varsigma(B \times [0,t])$ is $(\mathcal{F}'_t)$-adapted for all $B \in \mathcal{B}(\mathbb{R}^{d_1})$.

Given an adapted (in the above sense) $\mathcal{R}_1$-valued random variable ς, which corresponds to the control $u \in \mathcal{U}$, and a $\nu \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$, we will consider the controlled SDE

$$\begin{aligned} dX_t &= b(t,X,\nu_t)dt + \left(\int_{\mathbb{R}^{d_1}} \sigma(t,X,\nu_t) y \varsigma_t(dy) \right) dt + \sigma(t,X,\nu_t)dW_t \\ X(0) &= X_0, \end{aligned} \tag{1.13}$$

where W is a d_1-dimensional $(\mathcal{F}'_t)$-adapted standard Wiener process.
We deal with *weak solutions* of (1.13) or, equivalently, with certain probability measures on $\mathcal{B}(\mathcal{Z})$, where

$$\mathcal{Z} = \mathcal{C}([0,T], \mathbb{R}^d) \times \mathcal{R}_1 \times \mathcal{C}([0,T], \mathbb{R}^{d_1}).$$

For a typical element in $\mathcal{Z}$ let us write (φ, r, w) with the understanding that $\varphi \in \mathcal{C}([0,T], \mathbb{R}^d)$, $r \in \mathcal{R}_1$, $w \in \mathcal{C}([0,T], \mathbb{R}^{d_1})$.
The inclusion of W as a component of our canonical space $\mathcal{Z}$ will allow identification of the joint distribution of the control and driving Wiener process. Indeed, if the triple (X, ς, W) defined on some filtered probability space $(\Omega', \mathcal{F}', \mathbb{P}', (\mathcal{F}'_t))$ solves (1.13) for some continuous $\nu : [0,T] \to \mathcal{P}_1(\mathbb{R}^d)$, then the distribution of (X, ς, W) under $\mathbb{P}'$ is an element of $\mathcal{P}(\mathcal{Z})$.
Hence, $\mathcal{P}(\mathcal{Z})$ is the space we will focus on.

In fact, the question whether a probability measure $\Theta \in \mathcal{P}(\mathcal{Z})$ corresponds to a weak solution of the Equation (1.13), for a fixed flow of measures μ, can be conveniently phrased in terms of an associated local martingale problem.
Given $f \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}^{d_1})$, define a real valued process $(M_f^\infty(t))_{t\in[0,T]}$ on the probability space $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}), \Theta)$ by

$$M_f^\infty(t,(\varphi,r,w)) = f(\varphi(t),w(t)) - f(\varphi(0),w(0)) - \int_0^t \int_{\mathbb{R}^{d_1}} \mathcal{A}_s^\infty(f)(\varphi,r,w(s)) r_s(dy)ds, \tag{1.14}$$

where for $s \in [0,T], \varphi \in \mathcal{C}([0,T],\mathbb{R}^d)$ and $y, z \in \mathbb{R}^{d_1}$,

$$\begin{aligned}\mathcal{A}_s^\infty(f)(\varphi, y, z) =& \langle b(s,\varphi,\mu_s) + \sigma(s,\varphi,\mu_s)y, \nabla_x f(\varphi(s), z)\rangle \\ &+ \frac{1}{2}\sum_{j,k=1}^{d} (\sigma\sigma^T)_{jk}(s,\varphi,\mu_s)\frac{\partial^2 f}{\partial x_j \partial x_k}(\varphi(s), z) \\ &+ \frac{1}{2}\sum_{l=1}^{d_1} \frac{\partial^2 f}{\partial z_l^2}(\varphi(s), z) \\ &+ \frac{1}{2}\sum_{k=1}^{d}\sum_{l=1}^{d_1} \sigma_{kl}(s,\varphi,\mu_s)\frac{\partial^2 f}{\partial x_k \partial z_l}(\varphi(s), z)\end{aligned} \tag{1.15}$$

The expression involving $\mathcal{A}_s^\infty(f)$ in (1.14) is integrated against time and the time derivative measures r_s of any relaxed control r. Since we may use $r(dy \times dt)$ in place of $r_s(dy)ds$, the measures r_s are not actually needed.

The key tool is a one-to-one correspondence between weak solutions of (1.13), with $\nu = \mu$, and a local martingale problem.

Proposition 1.32. *Let $\Theta \in \mathcal{P}(\mathcal{Z})$ be such that $\Theta(\{(\varphi, r, w) \in \mathcal{Z} : w(0) = 0\}) = 1$. Then Θ corresponds to a weak solution of (1.13), with $\nu = \mu$, if and only if M_f^∞ is a local martingale under Θ with respect to the canonical filtration $(\mathcal{G}_t)$ for all $f \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}^{d_1})$.*

Proof. As in [3], we refer to the proof of Proposition 5.4.6 in [10], page 315. There is no need to extend the probability space $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}), \Theta)$ even if the diffusion coefficient σ is degenerate. In fact, the canonical process on the sample space $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ includes a component which corresponds to the driving Wiener process. □

Remark 1.33. *The canonical filtration $(\mathcal{G}_t)$ in $\mathcal{B}(\mathcal{Z})$ is not necessarily $\Theta-$complete or right continuous, while in the literature solutions to SDEs are usually defined with respect to filtrations satisfying the usual conditions (i.e., being right continuous and containing all sets contained in a set of measure zero). This is not a problem. In fact, it is possible to show that the local martingale property of the processes M_f^∞ under Θ with respect to the canonical filtration $(\mathcal{G}_t)$ implies that the canonical process on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ solves (1.13), with $\nu = \mu$, under Θ, with respect to the filtration $(\mathcal{G}_{t+}^\Theta)$ which satisfies the usual conditions. See [3].*

Let us recall that we are supposing that

$$\mathbb{E}\Big[\int_0^T |u(t)|^2 dt\Big] = Q < \infty, \tag{1.16}$$

This implies, in particular, that $\int_0^T |u(t,\omega)|^2 dt < \infty$, for $\mathbb{P}$-almost all $\omega \in \Omega$. Since we are working with weak solutions, modifying u on a set of $\mathbb{P}$-measure zero has no impact on our proof. Thus we may assume $\int_0^T |u(t,\omega)| dt < \infty$, for all $\omega \in \Omega$.

Let's go back to our proof. In order to show the continuity of the functional $\Psi(\cdot)$, we need to prove that, given a sequence $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$ convergent w.r.t. ρ_1 metric to $\mu \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$, $(\Psi(\mu_n))_{n\in\mathbb{N}}$ must converge to $\Psi(\mu)$, as $n \to \infty$, w.r.t. ρ_1.

$\forall n \in \mathbb{N}$, we define $\Theta_n \in \mathcal{P}(\mathcal{Z})$ as the weak solution of equation

$$\begin{aligned} dX_t &= b(t, X, \mu_n(t))dt + \int_{\mathbb{R}^{d_1}} \sigma(t, X, \mu_n(t)) y \varsigma_t(dy) dt + \sigma(t, X, \mu_n(t)) dW_t \\ X(0) &= X_0, \end{aligned} \tag{1.17}$$

and define Θ_μ as the weak solution of equation

$$\begin{aligned} dX_t &= b(t, X, \mu(t))dt + \int_{\mathbb{R}^{d_1}} \sigma(t, X, \mu(t)) y \varsigma_t(dy) dt + \sigma(t, X, \mu(t)) dW_t \\ X(0) &= X_0, \end{aligned} \tag{1.18}$$

where ς is the relaxed control corresponding to $u \in \mathcal{U}$. We can notice that Θ_n (resp. Θ_μ) corresponds to the law of the process (X_n, ς, W) (resp. (X^μ, ς, W)), where $(X_n(t))_{t\in[0,T]}$ (resp. $(X^\mu(t))_{t\in[0,T]}$) is the unique strong solution of the SDE (1.17) (resp. (1.18)).

We will show that $\Theta_n \to_{n\to\infty} \Theta_\mu$ in $\mathcal{P}(\mathcal{Z})$.
By *Mapping Theorem* and the continuity of projections we get that the sequence of first marginals of (Θ_n) converges to the first marginal of Θ_μ in $\mathcal{P}(\mathcal{C}([0,T], \mathbb{R}^d))$. Again, by *Mapping Theorem* and the continuity of the projections, this implies that $\Psi(\mu_n) \to \Psi(\mu)$, in $\mathcal{C}([0,T], \mathcal{P}(\mathbb{R}^d))$. This last fact, together with the fact, that we will show, that the sequence $(X_n)_{n\in\mathbb{N}}$ has finite moments of first order convergent to the first moment of the solution of the SDE (1.13) with $\nu = \mu$, implies that $\Psi(\mu_n) \to_{n\to\infty} \Psi(\mu)$, in $\mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$, w.r.t. ρ_1 metric, as desired.‖

So, once we have proved $\Theta_n \to_{n\to\infty} \Theta_\mu$ weakly, we can conclude.
First of all, we want to show that the sequence $(\Theta_n)_{n\in\mathbb{N}}$ is tight in $\mathcal{P}(\mathcal{Z})$. Since,

‖Here we have exploited the fact that Wasserstein metric, and so ρ_1, corresponds to the topology induced by weak convergence plus the convergence of first moments. See Theorem 7.12 in [15].

$\forall n \in \mathbb{N}$, Θ_n corresponds to the law of the process (X_n, ς, W), it is sufficient to prove tightness of the sequence $(\text{Law}(X_n))_{n\in\mathbb{N}}$ in $\mathcal{P}(\mathcal{C}([0,T],\mathbb{R}^d))$, in order to prove the tightness of (Θ_n). In order to prove that this sequence of laws is tight we will exploit *Aldous's tightness criterion*(Theorem 16.10 in [2]).

Theorem 1.34 (Aldous's tightness criterion)**.** *Consider a sequence of* $\mathcal{C}([0,T],\mathbb{R}^d)$*-valued random variables* $(X_n)_{n\in\mathbb{N}}$*. If the sequence satisfies the two conditions that follows then the sequence of their laws is tight.*

i)

$$\lim_{a\to\infty} \limsup_{n} \mathbb{P}\Big(\sup_{s\in[0,T]} |X_n(s)| \geq a \Big) = 0 \tag{1.19}$$

ii) For each $\epsilon, \eta > 0$*, there exist a* $\delta_0 > 0$ *and a* $n_0 \in \mathbb{N}$ *such that, if* $\delta \leq \delta_0$ *and* $n \geq n_0$ *and if* τ *is a discrete* X_n-*stopping time such that* $\tau \leq T$*, then*

$$\mathbb{P}(|X_n((\tau+\delta)\wedge T) - X_n(\tau)| \geq \epsilon) \leq \eta. \tag{1.20}$$

We can start by proving that condition i) holds.
Fix $a \in \mathbb{R}$. Exploiting (U) hypothesis and Markov inequality, we have $\forall n \in \mathbb{N}$:

$$\mathbb{P}\Big(\sup_{s\in[0,T]} |X_n(s)| \geq a \Big) \leq \frac{\mathbb{E}\Big[\sup_{s\in[0,T]} |X_n(s)|^2 \Big]}{a^2} \leq \frac{R'}{a^2}.$$

This estimation does not depend on n and let us deduce that

$$0 \leq \lim_{a\to\infty} \limsup_{n} \mathbb{P}\Big(\sup_{s\in[0,T]} |X_n(s)| \geq a \Big) \leq \lim_{a\to\infty} \frac{R'}{a^2} = 0,$$

that ends our proof on condition i).
Now, we have to prove condition ii). Fix arbitrarily $\epsilon, \eta > 0$ and let τ be a discrete X_n- stopping time such that $\tau \leq T$. Exploiting triangle inequality,

Hölder inequality, Itô isometry and Markov inequality, we get:

$$
\begin{aligned}
&\mathbb{P}(|X_n((\tau+\delta)\wedge T) - X_n(\tau)| \geq \epsilon) \\
&= \mathbb{P}\Big(\Big| \int_\tau^{(\tau+\delta)\wedge T} b(t, X_n, \mu_n(t))dt + \int_\tau^{(\tau+\delta)\wedge T} \sigma(t, X_n, \mu_n(t))u(t)dt \\
&\quad + \int_\tau^{(\tau+\delta)\wedge T} \sigma(t, X_n, \mu_n(t))dW_t\Big| \geq \epsilon\Big) \\
&\leq \mathbb{P}\Big(\int_\tau^{(\tau+\delta)\wedge T} |b(t, X_n, \mu_n(t))|dt + \Big| \int_\tau^{(\tau+\delta)\wedge T} \sigma(t, X_n, \mu_n(t))u(t)dt\Big| \\
&\quad + \Big| \int_\tau^{(\tau+\delta)\wedge T} \sigma(t, X_n, \mu_n(t))dW_t\Big| \geq \epsilon\Big) \\
&\leq \mathbb{P}\Big(\int_\tau^{(\tau+\delta)\wedge T} |b(t, X_n, \mu_n(t))|dt + \int_\tau^{(\tau+\delta)\wedge T} |\sigma(t, X_n, \mu_n(t))|^2 dt^{\frac{1}{2}} \\
&\quad \times \int_\tau^{(\tau+\delta)\wedge T} |u(t)|^2 dt^{\frac{1}{2}} \geq \frac{\epsilon}{2}\Big) + \mathbb{P}\Big(\Big| \int_\tau^{(\tau+\delta)\wedge T} \sigma(t, X_n, \mu_n(t))dW_t\Big| \geq \frac{\epsilon}{2}\Big) \\
&\leq \mathbb{P}\Big(\int_\tau^{(\tau+\delta)\wedge T} K(1 + \sup_{s\in[0,t]} |X_n(s)|)dt + K\sqrt{\delta}\Big(\int_\tau^{(\tau+\delta)\wedge T} |u(t)|^2 dt\Big)^{\frac{1}{2}} \geq \frac{\epsilon}{2}\Big) \\
&\quad + 4\frac{\mathbb{E}\Big[\int_\tau^{(\tau+\delta)\wedge T} |\sigma(t, X_n, \mu_n(t))|^2 d_t\Big]}{\epsilon^2} \\
&\leq \mathbb{P}\Big(\int_\tau^{(\tau+\delta)\wedge T} K(1 + \sup_{s\in[0,T]} |X_n(s)|)dt + K\sqrt{\delta}\Big(\int_0^T |u(t)|^2 dt\Big)^{\frac{1}{2}} \geq \frac{\epsilon}{2}\Big) + 4\frac{K^2\delta}{\epsilon^2} \\
&\leq \mathbb{P}\Big(\delta K(1 + \sup_{s\in[0,T]} |X_n(s)|) + K\sqrt{\delta}\Big(\int_0^T |u(t)|^2 dt\Big)^{\frac{1}{2}} \geq \frac{\epsilon}{2}\Big) + 4\frac{K^2\delta}{\epsilon^2} \\
&\leq 2\frac{\mathbb{E}\Big[\delta K(1 + \sup_{s\in[0,T]} |X_n(s)|)\Big] + K\sqrt{\delta}\mathbb{E}\Big[\int_0^T |u(t)|^2 dt\Big]^{\frac{1}{2}}}{\epsilon} + 4\frac{K^2\delta}{\epsilon^2} \\
&\leq 2\frac{K(1 + \sqrt{R'})\delta + K\sqrt{\delta}\sqrt{Q}}{\epsilon} + 4\frac{K^2\delta}{\epsilon^2} \\
&\leq 2\frac{K(1 + \sqrt{R'})\delta_0 + K\sqrt{\delta_0}\sqrt{Q}}{\epsilon} + 4\frac{K^2\delta_0}{\epsilon^2},
\end{aligned}
$$

which is clearly less than η with a good choice of δ_0, small enough.
Hence, $(\mathrm{Law}(X_n))_{n\in\mathbb{N}}$ is tight.

We have shown that $(\Theta_n)_{n\in\mathbb{N}} \subset \mathcal{P}(\mathcal{Z})$, $\Theta_n = \mathrm{Law}((X_n, \varsigma, W))$ $\forall n \in \mathbb{N}$, is tight.

By *Prokhorov's Theorem*, we know that any subsequence of (Θ_n) possesses

convergent sub-subsequence.
In the next lemma we identify the limit points of $(\Theta_n)_{n\in\mathbb{N}}$ as being weak solutions of Equation (1.18). By trajectorial uniqueness of solutions for Equation (1.13) for a fixed flux of measures θ, they must equal Θ_μ. This implies that $\Theta_n \to \Theta_\mu$, in $\mathcal{P}(\mathcal{Z})$, as $n \to \infty$.

Lemma 1.35. *Let $(\Theta_{N_j})_{j\in\mathbb{N}}$ be a weakly convergent subsequence of $(\Theta_n)_{n\in\mathbb{N}}$. Let Ξ be an element of $\mathcal{P}(\mathcal{Z})$ such that $\Theta_{N_j} \to \Xi$, as $j \to \infty$, weakly. Then Ξ is a weak solution of Equation (1.18). Hence, $\Xi = \Theta_\mu$.*

Proof. Set $I = \{N_j, j \in \mathbb{N}\}$, and write $(\Theta_n)_{n\in I}$ for $(\Theta_{N_j})_{j\in\mathbb{N}}$. By hypothesis $\Theta_n \to \Xi$, weakly. Recall from Proposition 1.32 that a probability measure $\Theta \in \mathcal{P}(\mathcal{Z})$ with $\Theta(\{(\phi, r, w) \in \mathcal{Z} : w(0) = 0\}) = 1$, corresponds to a weak solution of (1.13), with $\theta = \mu$, if (and only if), for all $f \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}^{d_1})$, M_f^∞ is a local martingale under Θ with respect to the canonical filtration $(\mathcal{G}_t)$, where M_f^∞ is defined in (1.14).
In verifying the local martingale property of M_f^∞ when $\Theta = \Xi$, we will work with randomized stopping times. Those stopping times live on an extension, $(\hat{\mathcal{Z}}, \mathcal{B}(\hat{\mathcal{Z}}))$, of the measurable space $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ and are adapted to a filtration $(\hat{\mathcal{G}}_t)$ in $\mathcal{B}(\hat{\mathcal{Z}})$, where

$$\hat{\mathcal{Z}} := \mathcal{Z} \times [0, 1], \quad \hat{\mathcal{G}}_t := \mathcal{G}_t \times \mathcal{B}([0, 1]), \quad t \in [0, T],$$

and $(\mathcal{G}_t)$ is the canonical filtration in $\mathcal{B}(\mathcal{Z})$. Any random object defined on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ also lives on $(\hat{\mathcal{Z}}, \mathcal{B}(\hat{\mathcal{Z}}))$, and no notational distinction will be made. Let λ denote the uniform distribution on $\mathcal{B}([0, 1])$. Any probability measure Θ on $\mathcal{B}(\mathcal{Z})$ induces a probability measure on $\mathcal{B}(\hat{\mathcal{Z}})$ given by $\hat{\Theta} := \Theta \times \lambda$. For each $k \in \mathbb{N}$, define a stopping time τ_k on $(\hat{\mathcal{Z}}, \mathcal{B}(\hat{\mathcal{Z}}))$, with respect to the filtration $(\hat{\mathcal{G}}_t)$ by setting for $(z, a) \in \mathcal{Z} \times [0, 1]$,

$$\tau_k(z, a) := \inf\{t \in [0, T] : v(t, z) \geq k + a\} \wedge T,$$

where

$$v((\varphi, r, w), t) := \int_{\mathbb{R}^{d_1}\times[0,t]} |y| r(dy \times ds) + \sup_{s\in[0,t]} |\varphi(s)| + \sup_{s\in[0,t]} |w(s)|.$$

Note that the mapping that associates to $t \in [0, T]$ $v((\varphi, r, w), t)$ is monotonic for all $(\varphi, r, w) \in \mathcal{Z}$. Hence, the stopping times have the following properties. The boundedness of φ and w (being continuous functions on a compact interval) and the boundedness of $\int_{\mathbb{R}^{d_1}\times[0,T]} |y| r(dy \times ds)$ imply that $\tau_k \to T$ as $k \to \infty$. The second property of note is that the mapping

$$\mathcal{Z} \times [0, 1] \ni (z, a) \mapsto \tau_k(z, a) \in [0, T]$$

is continuous with probability one under $\hat{\Theta} = \Theta \times \lambda$, $\forall \Theta \in \mathcal{P}(\mathcal{Z})$. To see this, note that for every $z \in \mathcal{Z}$ the set

$$A_z := \{c \in \mathbb{R}_+ : v(z,s) = c \text{ for all } s \in [t, t+\delta], \text{ some } t \in [0,T], \text{ some } \delta > 0\}$$

is at most countable. However, $\hat{z} \mapsto \tau_k(\hat{z})$ fails to be continuous at (z,a) only when $k + a \in A_z$. Therefore, by Fubini's Theorem,

$$\begin{aligned}
&\hat{\Theta}(\{(z,a) \in \hat{\mathcal{Z}} : \tau_k \text{ discontinuous at } (z,a)\}) \\
&= \int_{\hat{\mathcal{Z}}} \mathbb{I}_{A_z}(k+a)\hat{\Theta}(dz \times da) = \int_{\mathcal{Z}} \int_{[0,1]} \mathbb{I}_{A_z}(k+a)\lambda(da)\Theta(dz) = 0
\end{aligned}$$

Notice that if M_f^∞ is a local martingale with respect to $(\hat{\mathcal{G}}_t)$ under $\hat{\Theta} = \Theta \times \lambda$ with localizing sequence of stopping times $(\tau_k)_{k\in\mathbb{N}}$, then M_f^∞ is also a local martingale with respect to $(\mathcal{G}_t)$ under Θ with localizing sequence of stopping times $(\tau_k(\cdot, 0))_{k\in\mathbb{N}}$, see the Appendix in [3]. Thus, it suffices to prove the martingale property of M_f^∞ up till time τ_k with respect to filtration $(\hat{\mathcal{G}}_t)$ and probability measure $\hat{\Theta}$.
Clearly, the process $M_f^\infty(\cdot \wedge \tau_k)$ is a $(\hat{\mathcal{G}}_t)$-martingale under $\hat{\Theta}$ if and only if

$$\mathbb{E}_{\Theta\times\lambda}\Big[\psi \cdot (M_f^\infty(t_1 \wedge \tau_k) - M_f^\infty(t_0 \wedge \tau_k))\Big] = 0 \tag{1.21}$$

for all $t_0, t_1 \in [0,T]$ with $t_0 \leq t_1$, and $\hat{\mathcal{G}}_{t_0}$-measurable $\psi \in \mathcal{C}_b(\hat{\mathcal{Z}})$.

Let $(k, t_0, t_1, \psi, f) \in \mathbb{N} \times [0,T]^2 \times \mathcal{C}_b(\hat{\mathcal{Z}}) \times \mathcal{C}(\mathbb{R}^d \times \mathbb{R}^{d_1})$. For all $n \in \mathbb{N}$, Θ_n is a weak solution of Equation (1.13), with $\nu = \mu_n$. Proposition 1.32 implies that, $\forall f \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}^{d_1})$, $M_f^{(n)}$ is a local martingale on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}), \Theta_n)$, where $M_f^{(n)}$ is defined by:

$$M_f^{(n)}(t,(\varphi,r,w)) = f(\varphi(t),w(t)) - f(\varphi(0),w(0)) - \int_0^t \int_{\mathbb{R}^{d_1}} \mathcal{A}_s^{(n)}(f)(\varphi,r,w(s))r_s(dy)ds, \tag{1.22}$$

where for $s \in [0,T], \varphi \in \mathcal{C}([0,T],\mathbb{R}^d)$ and $y, z \in \mathbb{R}^{d_1}$,

$$\begin{aligned}
\mathcal{A}_s^{(n)}(f)(\varphi,y,z) =& \langle b(s,\varphi,\mu_n(s)) + \sigma(s,\varphi,\mu_n(s))y, \nabla_x f(\varphi(s),z)\rangle \\
&+ \frac{1}{2}\sum_{j,k=1}^{d} (\sigma\sigma^T)_{jk}(s,\varphi,\mu_n(s)) \frac{\partial^2 f}{\partial x_j \partial x_k}(\varphi(s),z) \\
&+ \frac{1}{2}\sum_{l=1}^{d_1} \frac{\partial^2 f}{\partial z_l^2}(\varphi(s),z) \\
&+ \frac{1}{2}\sum_{k=1}^{d}\sum_{l=1}^{d_1} \sigma_{kl}(s,\varphi,\mu_n(s)) \frac{\partial^2 f}{\partial x_k \partial z_l}(\varphi(s),z).
\end{aligned} \tag{1.23}$$

Hence, it must be true that:

$$\mathbb{E}_{\Theta_n\times\lambda}\Big[\psi\cdot(M_f^{(n)}(t_1\wedge\tau_k)-M_f^{(n)}(t_0\wedge\tau_k))\Big]=0. \tag{1.24}$$

As a consequence of assumption (UC) and by construction of the stopping time τ_k, the integrand in (1.21) is bounded; thanks to assumption (UC) and the almost sure continuity of τ_k, it is continuous with probability one under $\hat{\Xi}=\Xi\times\lambda$. By weak convergence and the *Mapping Theorem*, for instance Theorem 2.7 in [2], it follows that

$$\begin{aligned}\mathbb{E}_{\Theta_n\times\lambda}&\Big[\psi\cdot(M_f^{\infty}(t_1\wedge\tau_k)-M_f^{\infty}(t_0\wedge\tau_k))\Big]\\&\longrightarrow_{n\to\infty}\mathbb{E}_{\Xi\times\lambda}\Big[\psi\cdot(M_f^{\infty}(t_1\wedge\tau_k)-M_f^{\infty}(t_0\wedge\tau_k))\Big].\end{aligned}\tag{1.25}$$

Recall that, by hypothesis, we have $\mu_n\to\mu$ in ρ_1 metric.

We claim that this fact, together with the hypothesis (UC) and the construction of τ_k, implies that

$$\sup_{t\in[0,T]}\sup_{\hat{z}\in\hat{\mathcal{Z}}}\Big|M_f^{(n)}(t\wedge\tau_k(\hat{z}),\hat{z})-|M_f^{\infty}(t\wedge\tau_k(\hat{z}),\hat{z})|\Big|\longrightarrow_{n\to\infty}0.$$

In fact, if we consider, for example the integral, corresponding to the first term in the drift , which is

$$\int_0^{t\wedge\tau_k(\hat{z})}\langle b(s,\varphi,\mu_n(s)),\nabla_x f(\varphi(s),w(s))\rangle ds.$$

By the assumed uniform continuity properties of b, this converges uniformly in $t\in[0,T]$, $\hat{z}\in\hat{\mathcal{Z}}$ to

$$\int_0^{t\wedge\tau_k(\hat{z})}\langle b(s,\varphi,\mu(s)),\nabla_x f(\varphi(s),w(s))\rangle ds.$$

We can prove a similar result for all the other terms. Since ψ is bounded, it follows that

$$\begin{aligned}\Big|\mathbb{E}_{\Theta_n\times\lambda}&\Big[\psi\cdot(M_f^{\infty}(t_1\wedge\tau_k)-M_f^{\infty}(t_0\wedge\tau_k))\Big]\\&-\mathbb{E}_{\Theta_n\times\lambda}\Big[\psi\cdot(M_f^{(n)}(t_1\wedge\tau_k)-M_f^{(n)}(t_0\wedge\tau_k))\Big]\Big|\to_{n\to\infty}0.\end{aligned}\tag{1.26}$$

Remembering (1.24) and exploiting results (1.25) and (1.26) and triangle inequality , we have

$$\begin{aligned}
&\Big|\mathbb{E}_{\Xi\times\lambda}\Big[\psi\cdot(M_f^\infty(t_1\wedge\tau_k)-M_f^\infty(t_0\wedge\tau_k))\Big]\Big| \\
&=\Big|\mathbb{E}_{\Xi\times\lambda}\Big[\psi\cdot(M_f^\infty(t_1\wedge\tau_k)-M_f^\infty(t_0\wedge\tau_k))\Big] \\
&\qquad-\mathbb{E}_{\Theta_n\times\lambda}\Big[\psi\cdot(M_f^{(n)}(t_1\wedge\tau_k)-M_f^{(n)}(t_0\wedge\tau_k))\Big]\Big| \\
&\leq\Big|\mathbb{E}_{\Xi\times\lambda}\Big[\psi\cdot(M_f^\infty(t_1\wedge\tau_k)-M_f^\infty(t_0\wedge\tau_k))\Big] \\
&\qquad-\mathbb{E}_{\Theta_n\times\lambda}\Big[\psi\cdot(M_f^\infty(t_1\wedge\tau_k)-M_f^\infty(t_0\wedge\tau_k))\Big]\Big| \\
&+\Big|\mathbb{E}_{\Theta_n\times\lambda}\Big[\psi\cdot(M_f^\infty(t_1\wedge\tau_k)-M_f^\infty(t_0\wedge\tau_k))\Big] \\
&\qquad-\mathbb{E}_{\Theta_n\times\lambda}\Big[\psi\cdot(M_f^{(n)}(t_1\wedge\tau_k)-M_f^{(n)}(t_0\wedge\tau_k))\Big]\Big|\longrightarrow_{n\to\infty}0.
\end{aligned}$$

We have showed that

$$\mathbb{E}_{\Xi\times\lambda}\Big[\psi\cdot(M_f^\infty(t_1\wedge\tau_k)-M_f^\infty(t_0\wedge\tau_k))\Big]=0.$$

Hence, $\Xi=\Theta_\mu$, unique weak solution of SDE (1.18).

□

Consider the sequence $(X_n)_{n\in\mathbb{N}}$, where, $\forall n\in\mathbb{N}$, X_n is the unique strong solution of Equation (1.17). This sequence takes values in the Polish space $\mathcal{C}([0,T],\mathbb{R}^d)$. We must check the uniform convergence of the moments of first order of X_n to the moment of first order of X^μ, unique strong solution of the SDE (1.18). Namely, we should check that :

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[|X_n(t)|]-\mathbb{E}[|X^\mu(t)|]\Big|\longrightarrow_{n\to\infty}0.$$

We have already showed that the two following statements hold:

i) $X_n\longrightarrow X^\mu$, in distribution.

ii) $\sup\Big\{\sup_{n\in\mathbb{N}}\mathbb{E}\Big[\sup_{t\in[0,T]}|X_n(t)|^2\Big],\mathbb{E}\Big[\sup_{t\in[0,T]}|X^\mu(t)|^2\Big]\Big\}\leq R'<\infty.$

We claim that i) and ii) imply that

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[|X_n(t)|]-\mathbb{E}[|X^\mu(t)|]\Big|\longrightarrow_{n\to\infty}0.$$

Exploiting triangle inequality, we can write

$$\begin{aligned}
&\sup_{t\in[0,T]}\Big|\mathbb{E}[|X_n(t)|]-\mathbb{E}[|X^\mu(t)|]\Big| \\
&\leq\sup_{t\in[0,T]}\Big|\mathbb{E}\Big[|X_n(t)|\cdot\mathbb{I}_{|X_n(t)|\leq M}-|X^\mu(t)|\cdot\mathbb{I}_{|X^\mu(t)|\leq M}\Big]\Big| \\
&\qquad+\sup_{t\in[0,T]}\Big|\mathbb{E}\Big[|X_n(t)|\cdot\mathbb{I}_{|X_n(t)|>M}-|X^\mu(t)|\cdot\mathbb{I}_{|X^\mu(t)|>M}\Big]\Big|=A+B.
\end{aligned}$$

Exploiting Hölder inequality, Markov inequality and condition ii), we have:

$$B \leq \sup_{t\in[0,T]} \mathbb{E}\Big[|X_n(t)| \cdot \mathbb{I}_{|X_n(t)|>M}\Big] + \sup_{t\in[0,T]} \mathbb{E}\Big[|X^\mu(t)| \cdot \mathbb{I}_{|X^\mu(t)|>M}\Big]$$

$$\leq \sup_{t\in[0,T]} \mathbb{E}\Big[|X_n(t)|^2\Big]^{\frac{1}{2}} \mathbb{P}\Big(|X_n(t)| > M\Big)^{\frac{1}{2}} + \sup_{t\in[0,T]} \mathbb{E}\Big[|X^\mu(t)|^2\Big]^{\frac{1}{2}} \mathbb{P}\Big(|X^\mu(t)| > M\Big)^{\frac{1}{2}}$$

$$\leq \sqrt{R'}\Big(\sup_{t\in[0,T]} \mathbb{P}\Big(|X_n(t)| > M\Big)^{\frac{1}{2}} + \sup_{t\in[0,T]} \mathbb{P}\Big(|X^\mu(t)| > M\Big)^{\frac{1}{2}}\Big) \leq 2\frac{R'}{M},$$

which can be made smaller than $\frac{\epsilon}{3}$ for a suitable choice of $M \geq \frac{6R'}{\epsilon}$.
Analogously, we get:

$$A = \sup_{t\in[0,T]} \Big|\mathbb{E}\Big[|X_n(t)| \wedge M \cdot \mathbb{I}_{|X_n(t)|\leq M} - |X^\mu(t)| \wedge M \cdot \mathbb{I}_{|X^\mu(t)|\leq M}\Big]\Big|$$

$$\leq \sup_{t\in[0,T]} \Big|\mathbb{E}\Big[|X_n(t)| \wedge M - |X^\mu(t)| \wedge M\Big]\Big|$$

$$+ \sup_{t\in[0,T]} \Big|\mathbb{E}\Big[|X_n(t)| \wedge M \cdot \mathbb{I}_{|X_n(t)|>M} - |X^\mu(t)| \wedge M \cdot \mathbb{I}_{|X^\mu(t)|>M}\Big]\Big|$$

$$\leq \sup_{t\in[0,T]} \Big|\mathbb{E}\Big[|X_n(t)| \wedge M - |X^\mu(t)| \wedge M\Big]\Big|$$

$$+ M\Big(\sup_{t\in[0,T]} \mathbb{P}\Big(|X_n(t)| > M\Big) + \sup_{t\in[0,T]} \mathbb{P}\Big(|X^\mu(t)| > M\Big)\Big)$$

$$\leq \sup_{t\in[0,T]} \Big|\mathbb{E}\Big[|X_n(t)| \wedge M - |X^\mu(t)| \wedge M\Big]\Big| + 2M\frac{R'}{M^2}$$

As above, $2M\frac{R'}{M^2} = \frac{2R'}{M}$ can be made smaller than $\frac{\epsilon}{3}$ choosing M big enough. Furthermore, hypothesis i) guarantees that, for every M fixed, $\sup_{t\in[0,T]} \Big|\mathbb{E}\Big[|X_n(t)| \wedge M - |X^\mu(t)| \wedge M\Big]\Big| \leq \frac{\epsilon}{3}$ for $n \geq n_M$.
Hence, we have

$$\sup_{t\in[0,T]} \Big|\mathbb{E}[|X_n(t)|] - \mathbb{E}[|X^\mu(t)|]\Big| \leq \epsilon, \forall n \geq n_M,$$

which let us conclude by the arbitrariness of $\epsilon > 0$.

This ends our proof of the continuity of functional $\Psi(\cdot)$ and proves the Existence of solutions for MV-SDEs under relaxed hypotheses.

Remark 1.36. *We would like to underline that the hypothesis of uniqueness of solutions in (U) is not strictly necessary. Removing that condition we will get a $\Psi(\cdot)$ which is no more a function but a correspondence. In this case to show Existence of solutions for our SDEs we could exploit Theorem 1.17 and limit points must be included in the set of weak solutions for SDE (1.18).*

Chapter 2

Uniqueness of solutions

2.1 Control with finite exponential moments

Let $T > 0$ be a time horizon and $d, d_1 \in \mathbb{N}$. Equip $\mathcal{C}([0,T], \mathbb{R}^d)$ with the supremum-norm topology. Let $\mathcal{P}(\mathbb{R}^d)$ be the space of probability measures on $\mathbb{R}^d$.
Let $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t\in[0,T]})$ be a stochastic basis satisfying the usual hypotheses and carrying a d_1-dimensional $(\mathcal{F})_t$-Wiener process $(W_t)_{t\in[0,T]}$.
Let d_{bL} be the bounded Lipschitz metric on $\mathcal{P}(\mathbb{R}^d)$, that is, given $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$,

$$d_{bL}(\mu, \nu) = \sup \left\{ \int_{\mathbb{R}^d} f(x)\mu(dx) - \int_{\mathbb{R}^d} f(x)\nu(dx) : \|f\|_{bL} \leq 1 \right\},$$

where $\|.\|_{bL}$ is defined for bounded Lipschitz functions $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ by

$$\|f\|_{bL} = \sup_{x\in\mathbb{R}^d} |f(x)| + \sup_{x\neq y\in\mathbb{R}^d} \frac{|f(x) - f(y)|}{|x-y|}.$$

Remark 2.1. *Let us recall that this metric induces the topology of weak convergence. See, for instance, chapter 11.3 in [6].*

It possible to show the following estimations.

Theorem 2.2. *If X, Y are two $\mathbb{R}^d$-valued random variables on the same probability space and A is in $\mathcal{F}$, both the following are true:*

$$d_{bL}(Law(X), Law(Y)) \leq \mathbb{E}[|X - Y|], \tag{2.1}$$

$$d_{bL}(Law(X), Law(Y)) \leq \mathbb{E}[|X - Y| \cdot \mathbb{I}_A] + 2\mathbb{P}(A^C). \tag{2.2}$$

Proof. Let X, Y be two $\mathbb{R}^d$-valued random variables on the same probability space. Consider $f : \mathbb{R}^d \longrightarrow \mathbb{R}$, such that $\|f\|_{bL} \leq 1$. This means that, $\forall x, y \in$

$\mathbb{R}^d$, we have:

$$|f(x)-f(y)| = \frac{|f(x)-f(y)|}{|x-y|}|x-y| \le \sup_{v\neq w}\frac{|f(v)-f(w)|}{|v-w|}|x-y| \le \|f\|_{bL}|x-y| \\ \le |x-y|.$$

We are now able to show the first result:

$$d_{bL}(\mathrm{Law}(X),\mathrm{Law}(Y)) \\ = \sup\Big\{\int_{\mathbb{R}^d} f(x)\mathrm{Law}(X)(dx) - \int_{\mathbb{R}^d} f(x)\mathrm{Law}(Y)(dx) : \|f\|_{bL}\le 1\Big\} \\ = \sup\Big\{\mathbb{E}[f(X)]-\mathbb{E}[f(Y)] : \|f\|_{bL}\le 1\Big\} \le \sup\Big\{\mathbb{E}[|f(X)-f(Y)|] : \|f\|_{bL}\le 1\Big\} \\ \le \mathbb{E}\Big[|X-Y|\Big].$$

In order to prove Property (2.2), consider $A\in\mathcal{F}$, we have that:

$$d_{bL}(\mathrm{Law}(X),\mathrm{Law}(Y)) = \sup\Big\{\mathbb{E}[f(X)]-\mathbb{E}[f(Y)] : \|f\|_{bL}\le 1\Big\} \\ \le \sup\Big\{\mathbb{E}[|f(X)-f(Y)|] : \|f\|_{bL}\le 1\Big\} \\ \le \sup\Big\{\mathbb{E}[|f(X)-f(Y)|\cdot\mathbb{I}_A] + \mathbb{E}[|f(X)-f(Y)|\cdot\mathbb{I}_{A^C}] : \|f\|_{bL}\le 1\Big\} \\ \le \sup\Big\{\mathbb{E}[|f(X)-f(Y)|\cdot\mathbb{I}_A] + \mathbb{E}[2\|f\|_\infty\mathbb{I}_{A^C}] : \|f\|_{bL}\le 1\Big\} \\ \le \sup\Big\{\mathbb{E}[|X-Y|\cdot\mathbb{I}_A] + 2\|f\|_{bL}\mathbb{P}(A^C) : \|f\|_{bL}\le 1\Big\} \\ \le \mathbb{E}\Big[|X-Y|\mathbb{I}_A\Big] + 2\mathbb{P}(A^C).$$

This ends our proof. □

Let b,σ be predictable functionals on $[0,T]\times\mathcal{C}([0,T],\mathbb{R}^d)\times\mathcal{P}(\mathbb{R}^d)$ with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d_1}$ respectively.
Now, consider the following non-linear SDE:

$$dX_t = b(t,X,\mathrm{Law}(X(t)))dt + \sigma(t,X,\mathrm{Law}(X(t)))u(t)dt \\ + \sigma(t,X,\mathrm{Law}(X(t)))dW_t, \tag{2.3}$$

with $X(0)=X_0$ fixed, such that $\mathbb{E}\big[|X_0|^2\big]<\infty$, and u a $\mathbb{R}^{d_1}$-valued $(\mathcal{F}_t)$-progressively measurable process such that

$$\mathbb{E}\left[\int_0^T |u(t)|^2 dt\right] < \infty.$$

Consider the following Lipschitz and growth conditions on b and σ.

(L) There exists L such that for all $t \in [0,T]$, all $\phi, \psi \in \mathcal{C}([0,T],\mathbb{R}^d)$ all $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$

$$|b(t,\phi,\mu)-b(t,\psi,\nu)|+|\sigma(t,\phi,\mu)-\sigma(t,\psi,\nu)| \leq L\Big(\sup_{s\in[0,t]}|\phi(s)-\psi(s)|+d_{bL}(\mu,\nu)\Big).$$

(G) There exists a constant $K > 0$ such that for all $t \in [0,T]$, all $\phi \in \mathcal{C}([0,T],\mathbb{R}^d)$, all $\mu \in \mathcal{P}(\mathbb{R}^d)$

$$|b(t,\phi,\mu)| \leq K\Big(1+\sup_{s\in[0,t]}|\phi(s)|\Big), \qquad |\sigma(t,\phi,\mu)| \leq K.$$

Let us notice that these conditions are sufficient to guarantee the finiteness of the moments of second order of the solutions of the SDE.*

Proposition 2.3. *Grant conditions (L) and (G). Let* $((\Omega,\mathcal{F},\mathbb{P}),(\mathcal{F}_t)_{t\in[0,T]})$ *be a stochastic basis satisfying the usual hypotheses and carrying a* d_1*-dimensional* $(\mathcal{F})_t$*-Wiener process* $(W_t)_{t\in[0,T]}$ *and let* u *be a* $\mathbb{R}^{d_1}$*-valued* $(\mathcal{F}_t)$*-progressively measurable process such that*

$$\mathbb{E}\left[\int_0^T |u(t)|^2 dt\right] = Q < \infty.$$

Suppose that $X, \tilde{X}$ *are solutions of (2.3) over the time interval* $[0,T]$*, under the control* u*, with initial condition* $X(0) = X_0 = \tilde{X}(0)$*,* $\mathbb{P}$*-almost surely.*
Then, for all $M \in \mathbb{N}$*, we have:*

$$d_{bL}(Law(X(t)), Law(\tilde{X}(t)))^2 \leq \frac{8Q(t)^2}{M^2} + 2K_M\int_0^t e^{2K_M(t-s)}\frac{8Q(s)^2}{M^2}ds,$$

where $Q(t) = \mathbb{E}\Big[\int_0^t |u(s)|^2 ds\Big] \leq \mathbb{E}\Big[\int_0^T |u(s)|^2 ds\Big] = Q < \infty$ *and*
$K_M = B_M + (B_M)^2 e^{B_M T} T$*, with* $B_M = 8L^2(T+M+4)$*.*

Proof. For $M \in \mathbb{N}$, define a $(\mathcal{F}_t)$-stopping time τ_M by:

$$\tau_M(\omega) = \inf\{t \in [0,T] : \int_0^t |u(s,\omega)|^2 ds \geq M\},$$

with the obvious assumption $\inf(\varnothing) = \infty$. Observe that $\mathbb{P}(\tau_M \leq T) \to 0$ as $M \to \infty$, thanks to the fact that $\mathbb{E}\Big[\int_0^T |u(t)|^2 dt\Big] < \infty$.
Set $\theta(t) = \text{Law}(X(t)), \tilde{\theta}(t) = \text{Law}(\tilde{X}(t)), t \in [0,T]$.

*See Appendix B, in particular Remark B.3.

Using Hölder's inequality, Doob's maximal inequality, the Itô isometry and condition (L), we obtain for $M \in \mathbb{N}$, all $t \in [0,T]$,

$$\begin{aligned}
\mathbb{E}\left[\sup_{s\in[0,t]} |X(s\wedge\tau_M) - \tilde{X}(s\wedge\tau_M)|^2\right] &\le 4T\mathbb{E}\left[\int_0^{t\wedge\tau_M} |b(r,X,\theta(r)) - b(r,\tilde{X},\tilde{\theta}(r))|^2 dr\right] \\
&+ 4\mathbb{E}\left[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr \cdot \int_0^{t\wedge\tau_M} |u(r)|^2 dr\right] \\
&+ 16\mathbb{E}\left[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr\right] \\
&\le 4T\mathbb{E}\left[\int_0^{t\wedge\tau_M} |b(r,X,\theta(r)) - b(r,\tilde{X},\tilde{\theta}(r))|^2 dr\right] \\
&+ (4M+16)\mathbb{E}\left[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr\right] \\
&\le 8L^2(T+M+4)\mathbb{E}\left[\int_0^{t\wedge\tau_M} \Big(\sup_{s\in[0,r]} |X(s) - \tilde{X}(s)|^2 + d_{bL}(\theta(r),\tilde{\theta}(r))^2\Big) dr\right] \\
&\le 8L^2(T+M+4)\int_0^t \mathbb{E}[\sup_{s\in[0,r]} |X(s\wedge\tau_M) - \tilde{X}(s\wedge\tau_M)|^2] dr \\
&+ 8L^2(T+M+4)\int_0^t d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr.
\end{aligned}$$

We recall the following generalized version of Bellman-Gronwall's Lemma.

Lemma 2.4. *Let I denote an interval of the real line of the form $[a,\infty)$ or $[a,b]$ or $[a,b)$, with $a < b$. Let α, b and u be real-valued functions defined on I. Assume that b and u are continuous and that the negative part of α is integrable on every closed and bounded subinterval of I. Furthermore, assume that b is non-negative and u satisfies the integral inequality:*

$$u(t) \le \alpha(t) + \int_a^t b(s)u(s)ds, \qquad \forall t \in I.$$

Then

$$u(t) \le \alpha(t) + \int_a^t \alpha(s)b(s)\exp(\int_s^t b(r)dr)ds \qquad t \in I.$$

Applying that Lemma with

$$b(s) = +8L^2(T+M+4)(= B_M),$$
$$\alpha(s) = +8L^2(T+M+4)\int_0^s d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr = B_M \int_0^s d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr,$$

we can deduce the following estimation:

$$\mathbb{E}\left[\sup_{s\in[0,t]} |X(s\wedge\tau_M) - \tilde{X}(s\wedge\tau_M)|^2\right] \leq \alpha(t) + B_M \int_0^t \alpha(s)e^{B_M(t-s)}ds =$$

$$= B_M \int_0^t d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr + B_M \int_0^t B_M \int_0^s d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr e^{B_M(t-s)}ds$$

$$\leq K_M \int_0^t d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr,$$

with $K_M = B_M + (B_M)^2 e^{B_M T}T$.
Furthermore, using property (2.2), Cauchy-Schwarz inequality and the well-known fact that $(a+b)^2 \leq 2(a^2+b^2)$, we have that

$$d_{bL}(\theta(t),\tilde{\theta}(t))^2 \leq 2\mathbb{E}[|X(t)-\tilde{X}(t)|^2 \mathbb{I}_{t<\tau_M}] + 8\mathbb{P}(t\geq\tau_M)^2$$

$$\leq 2\mathbb{E}[|X(t\wedge\tau_M)-\tilde{X}(t\wedge\tau_M)|^2] + 8\mathbb{P}\left(\int_0^t |u(s)|^2 ds \geq M\right)^2.$$

Exploiting what we have shown above and Markov inequality, we have that

$$d_{bL}(\theta(t),\tilde{\theta}(t))^2 \leq 2K_M \int_0^t d_{bL}(\theta(s),\tilde{\theta}(s))^2 ds + 8\Big(\frac{\mathbb{E}(\int_0^t |u(s)|^2 ds)}{M}\Big)^2$$

$$\leq 2K_M \int_0^t d_{bL}(\theta(s),\tilde{\theta}(s))^2 ds + \frac{8Q(t)^2}{M^2},$$

where $Q(t) = \mathbb{E}(\int_0^t |u(s)|^2 ds)$.
Now applying Lemma (2.4) a second time, we get:

$$d_{bL}(\theta(t),\tilde{\theta}(t))^2 \leq \frac{8Q(t)^2}{M^2} + 2K_M \int_0^t e^{2K_M(t-s)} \frac{8Q(s)^2}{M^2} ds.$$

□

We get a stronger result if we improve our hypotheses.

Proposition 2.5. *Grant conditions (L) and (G). Let* $((\Omega,\mathcal{F},\mathbb{P}),(\mathcal{F}_t)_{t\in[0,T]})$ *be a stochastic basis satisfying the usual hypotheses and carrying a* d_1*-dimensional* $(\mathcal{F})_t$*-Wiener process* $(W_t)_{t\in[0,T]}$ *and let* u *be a* $\mathbb{R}^{d_1}$*-valued* $(\mathcal{F}_t)$*-progressively measurable process such that there exists a positive constant,* $c>0$*, such that :*

$$\mathbb{E}\left[e^{c\int_0^T |u(t)|^2 dt}\right] < \infty.$$

Suppose that $X, \tilde{X}$ *are solutions of (2.3), over the time interval* $[0,T]$*, under the control* u*, with initial condition* $X(0) = X_0 = \tilde{X}(0)$*,* $\mathbb{P}$*-almost surely. Then* $X, \tilde{X}$ *are indistinguishable, that is :*

$$\mathbb{P}(X(t) = \tilde{X}(t), \forall t\in[0,T]) = 1.$$

Proof. For $M \in \mathbb{N}$, define a $(\mathcal{F}_t)$-stopping time τ_M by:

$$\tau_M(\omega) = \inf\{t \in [0,T] : \int_0^t |u(s,\omega)|^2 ds \geq M\},$$

with the obvious assumption $\inf(\varnothing) = \infty$. Observe that $\mathbb{P}(\tau_M \leq T) \to 0$ as $M \to \infty$, thanks to the fact that $\mathbb{E}\Big[\int_0^T |u(t)|^2 dt\Big] < \infty$.
Set $\theta(t) = \mathrm{Law}(X(t)), \tilde{\theta}(t) = \mathrm{Law}(\tilde{X}(t)), t \in [0,T]$.
Choose $\delta \in \Big(0, \frac{c}{12L^2}\Big)$ such that $T = l\delta$, for some $l \in \mathbb{N}$. We get a finite covering of the compact interval $[0,T]$ of the type $\bigcup_{k=0}^{l-1}[k\delta,(k+1)\delta] = [0,T]$.
Consider the first subinterval of time $[0,\delta]$ and let $t \leq \delta$. Using Hölder's inequality, Doob's maximal inequality, the Itô isometry and condition (L), we obtain, for $M \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{E}\Big[\sup_{s\in[0,t]} |X(s\wedge\tau_M) - \tilde{X}(s\wedge\tau_M)|^2\Big] &\leq 4\delta\mathbb{E}\Big[\int_0^{t\wedge\tau_M} |b(r,X,\theta(r)) - b(r,\tilde{X},\tilde{\theta}(r))|^2 dr\Big] \\
&+ 4\mathbb{E}\Big[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr \cdot \int_0^{t\wedge\tau_M} |u(r)|^2 dr\Big] \\
&+ 16\mathbb{E}\Big[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr\Big] \\
&\leq 4\delta\mathbb{E}\Big[\int_0^{t\wedge\tau_M} |b(r,X,\theta(r)) - b(r,\tilde{X},\tilde{\theta}(r))|^2 dr\Big] \\
&+ (4M+16)\mathbb{E}\Big[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr\Big] \\
&\leq 8L^2(\delta+M+4)\mathbb{E}\Big[\int_0^{t\wedge\tau_M} \Big(\sup_{s\in[0,r]} |X(s) - \tilde{X}(s)|^2 + d_{bL}(\theta(r),\tilde{\theta}(r))^2\Big) dr\Big].
\end{aligned}$$

Exploiting the property (2.2) of $d_{bL}(\cdot,\cdot)$ and the fact that $(a+b)^2 \leq 2(a^2+b^2)$, we deduce:

$$\mathbb{E}\left[\sup_{s\in[0,t]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\right]$$
$$\leq 8L^2(\delta+M+4)\mathbb{E}\left[\int_0^t\sup_{s\in[0,r]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2dr\right]$$
$$+8L^2(\delta+M+4)\mathbb{E}\left[\int_0^{t\wedge\tau_M}\Big(\mathbb{E}[|X(r)-\tilde{X}(r)|\mathbb{I}_{t<\tau_M}]+2\mathbb{P}(t\geq\tau_M)\Big)^2dr\right]$$
$$=8L^2(\delta+M+4)\Bigg(\int_0^t\mathbb{E}[\sup_{s\in[0,r]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2]dr$$
$$+\mathbb{E}\left[\int_0^{t\wedge\tau_M}2\mathbb{E}\Big[|X(r\wedge\tau_M)-\tilde{X}(r\wedge\tau_M)|^2\Big]+8\mathbb{P}\Big(\int_0^t|u(v)|^2dv\geq M\Big)^2dr\right]\Bigg)$$
$$\leq 24L^2(\delta+M+4)\mathbb{E}\left[\int_0^t\sup_{s\in[0,r]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2dr\right]$$
$$+64L^2(\delta+M+4)\int_0^t\mathbb{P}\Big(\int_0^t|u(v)|^2dv\geq M\Big)^2dr$$
$$\leq 24L^2(\delta+M+4)\mathbb{E}\left[\int_0^t\sup_{s\in[0,r]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2dr\right]$$
$$+64L^2(\delta+M+4)\delta\mathbb{P}\Big(\int_0^T|u(v)|^2dv\geq M\Big)^2.$$

Exploiting the assumptions on the finiteness of the moments of exponential order of u together with the fact that
$\mathbb{P}\Big[\int_0^T|u(v)|^2dv\geq M\Big]=\mathbb{P}\Big[e^{c\int_0^T|u(v)|^2dv}\geq e^{cM}\Big]$,
using Markov inequality, we can rewrite the previous as:

$$\mathbb{E}\left[\sup_{s\in[0,t]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\right]$$
$$\leq 24L^2(\delta+M+4)\mathbb{E}\left[\int_0^t\sup_{s\in[0,r]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2dr\right]$$
$$+64L^2(\delta+M+4)\delta\frac{\mathbb{E}\Big[e^{c\int_0^T|u(v)|^2dv}\Big]^2}{e^{2cM}}$$
$$\leq 3B_M\mathbb{E}\left[\int_0^t\sup_{s\in[0,r]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2dr\right]+8B_M\delta\frac{V^2}{e^{2cM}},$$

where B_M as above and $V=\mathbb{E}\Big[e^{c\int_0^T|u(v)|^2dv}\Big]$.

Now, applying Gronwall Lemma to the previous, we get the estimation:

$$\mathbb{E}\left[\sup_{s\in[0,t]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\right]\leq 8B_M\delta\frac{V^2}{e^{2cM}}e^{3B_M t}.$$

Hence, for every $t\in[0,\delta]$ we have the following

$$\mathbb{E}\left[\sup_{s\in[0,t]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\right]\leq 8B_M\delta V^2 e^{-2cM+3B_M\delta}.$$

Since $c>12\delta L^2$ and $B_M=8L^2(\delta+M+4)$, at the limit for $M\to\infty$, applying the monotone convergence Theorem, we have:

$$\mathbb{E}\left[\sup_{s\in[0,\delta]}|X(s)-\tilde{X}(s)|^2\right]=\lim_{M\to\infty}\mathbb{E}\left[\sup_{s\in[0,\delta]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\right]$$
$$\leq\lim_{M\to\infty}8B_M\delta V^2 e^{-2cM+3B_M\delta}=0,$$

and so $\mathbb{P}(X(t)=\tilde{X}(t),0\leq t\leq\delta)=1$.
Now, we can repeat the same argument for the time interval $[0,2\delta]$, exploiting that we know that is $X(s)=\tilde{X}(s)\quad,\forall s\in[0,\delta],\quad\mathbb{P}-a.s.$.
Let us notice that, in particular, this implies that :

$$\sup_{s\in[0,t]}|X(s)-\tilde{X}(s)|=\sup_{s\in[\delta,t]}|X(s)-\tilde{X}(s)|,\quad\mathbb{P}-a.s.,\quad\text{for } t\in[\delta,2\delta],$$
$$\sup_{s\in[0,t]}|X(s)-\tilde{X}(s)|=0,\quad\mathbb{P}-a.s.,\quad\text{for } t\in[0,\delta].$$

Furthermore, we have that:

$$d_{bL}(\theta(t),\tilde{\theta(t)})=0,\qquad t\in[0,\delta].$$

For $M\in\mathbb{N}$, we recall the definition of the $(\mathcal{F}_t)$-stopping time τ_M:

$$\tau_M(\omega)=\inf\{t\in[0,T]:\int_0^t|u(s,\omega)|^2ds\geq M\}.$$

Let $t\in[\delta,2\delta]$.

Exploiting the definitions of X and $\tilde{X}$ and the triangle inequality, we have:

$$
\begin{aligned}
&\mathbb{E}\bigg[\sup_{s\in[0,t]} |X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\bigg] \\
&=\mathbb{E}\bigg[\sup_{s\in[0,t]} |\int_0^{s\wedge\tau_M}(\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r)))dW_r \\
&\quad+\int_0^{s\wedge\tau_M}(b(r,X,\theta(r))-b(r,\tilde{X},\tilde{\theta}(r)))+(\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r)))u(r)dr|^2\bigg] \\
&\le 4\mathbb{E}\bigg[\sup_{s\in[0,t]} |\int_0^{s\wedge\tau_M}|b(r,X,\theta(r))-b(r,\tilde{X},\tilde{\theta}(r))|dr|^2\bigg] \\
&\quad+4\mathbb{E}\bigg[\sup_{s\in[0,t]} |\int_0^{s\wedge\tau_M}|\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r))||u(r)|dr|^2\bigg] \\
&\quad+4\mathbb{E}\bigg[\sup_{s\in[0,t]} |\int_0^{s\wedge\tau_M}(\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r)))dW_r|^2\bigg]=\triangle.
\end{aligned}
$$

Applyng hypothesis (L) to the first term, Hölder's inequality to the second and Doob's maximal inequality, we can write:

$$
\begin{aligned}
\triangle &\le 4L^2\mathbb{E}\bigg[\sup_{s\in[0,t]} |\int_0^{s\wedge\tau_M}\sup_{w\in[0,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})dr|^2\bigg] \\
&\quad+4\mathbb{E}\bigg[\sup_{s\in[0,t]} \int_0^{s\wedge\tau_M}|\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r))|^2dr\cdot\int_0^{s\wedge\tau_M}|u(r)|^2dr\bigg] \\
&\quad+4\Big(\frac{2}{2-1}\Big)^2\mathbb{E}\bigg[|\int_0^{t\wedge\tau_M}(\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r)))dW_r|^2\bigg]=\triangledown.
\end{aligned}
$$

Exploiting the fact that $\int_0^{s\wedge\tau_M}|u(r)|^2dr\le M$ and Itô isometry, we have that:

$$
\begin{aligned}
\triangledown &\le 4L^2\mathbb{E}\bigg[\sup_{s\in[0,t]} |\int_0^{s\wedge\tau_M}\sup_{w\in[0,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})dr|^2\bigg] \\
&\quad+4M\mathbb{E}\bigg[\sup_{s\in[0,t]} \int_0^{s\wedge\tau_M}|\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r))|^2dr\bigg] \\
&\quad+16\mathbb{E}\bigg[\int_0^{t\wedge\tau_M}|\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r))|^2dr\bigg]=\bigcirc.
\end{aligned}
$$

Using the sup properties and hypothesis (L) once more, we can deduce:

$$\bigcirc \leq 4L^2\mathbb{E}\left[\sup_{s\in[0,t]}|\int_0^{s\wedge\tau_M}\sup_{w\in[0,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})dr|^2\right]$$
$$+(4M+16)\mathbb{E}\left[\int_0^{t\wedge\tau_M}|\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r))|^2dr\right]$$
$$\leq 4L^2\mathbb{E}\left[|\int_0^{t\wedge\tau_M}\sup_{w\in[0,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})dr|^2\right]$$
$$+(4M+16)\mathbb{E}\left[\int_0^{t\wedge\tau_M}|\sigma(r,X,\theta(r))-\sigma(r,\tilde{X},\tilde{\theta}(r))|^2dr\right]$$
$$\leq 4L^2\mathbb{E}\left[|\int_0^{t\wedge\tau_M}\sup_{w\in[0,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})dr|^2\right]$$
$$+(4M+16)L^2\mathbb{E}\left[\int_0^{t\wedge\tau_M}\sup_{w\in[0,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})|^2dr\right]=\Diamond.$$

Furthermore, remembering that

$$\sup_{s\in[0,t]}|X(s)-\tilde{X}(s)|=0,\qquad \text{for } t\in[0,\delta],\qquad \mathbb{P}-a.s.$$
$$\sup_{s\in[0,t]}|X(s)-\tilde{X}(s)|=\sup_{s\in[\delta,t]}|X(s)-\tilde{X}(s)|,\qquad \text{for } t\in[\delta,2\delta],\qquad \mathbb{P}-a.s.$$
$$d_{bL}(\theta(t),\tilde{\theta(t)})=0,\qquad t\in[0,\delta],$$

we can write:

$$\Diamond \leq 4L^2\mathbb{E}\left[|\int_{\delta\wedge\tau_M}^{t\wedge\tau_M}\sup_{w\in[\delta,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})dr|^2\right]$$
$$+(4M+16)L^2\mathbb{E}\left[\int_{\delta\wedge\tau_M}^{t\wedge\tau_M}|\sup_{w\in[\delta,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})|^2dr\right]=\spadesuit.$$

Exploiting Hölder inequality applied to the first term, we get:

$$\spadesuit \leq 4L^2\delta\mathbb{E}\left[\int_{\delta\wedge\tau_M}^{t\wedge\tau_M}\Big(\sup_{w\in[\delta,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})\Big)^2dr\right]$$
$$+(4M+16)L^2\mathbb{E}\left[\int_{\delta\wedge\tau_M}^{t\wedge\tau_M}\Big(\sup_{w\in[\delta,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})\Big)^2dr\right]$$
$$\leq 4L^2(\delta+M+4)\mathbb{E}\left[\int_{\delta\wedge\tau_M}^{t\wedge\tau_M}\Big(\sup_{w\in[\delta,r]}|X_w-\tilde{X}_w|+d_{bL}(\theta(r),\tilde{\theta(r)})\Big)^2dr\right]$$

$$\leq 8L^2(\delta+M+4)\mathbb{E}\Bigg[\int_{\delta\wedge\tau_M}^{t\wedge\tau_M}\sup_{w\in[\delta,r]}|X_w-\tilde{X}_w|^2+d_{bL}(\theta(r),\tilde{\theta(r)})^2dr\Bigg]$$
$$\leq 8L^2(\delta+M+4)\mathbb{E}\Bigg[\int_{\delta}^{t}\sup_{w\in[0,r]}|X_{w\wedge\tau_M}-\tilde{X}_{w\wedge\tau_M}|^2+d_{bL}(\theta(r),\tilde{\theta(r)})^2dr\Bigg]$$
$$\leq 8L^2(\delta+M+4)\Bigg(\int_{\delta}^{t}\mathbb{E}\Bigg[\sup_{w\in[0,r]}|X_{w\wedge\tau_M}-\tilde{X}_{w\wedge\tau_M}|^2\Bigg]dr+\int_{\delta}^{t}d_{bL}(\theta(r),\tilde{\theta(r)})^2dr\Bigg).$$

Exploiting the Property (2.2) of $d_{bL}(\cdot,\cdot)$, the fact that $(a+b)^2 \leq 2(a^2+b^2)$ and Cauchy-Schwarz inequality, we deduce:

$$\begin{aligned}
&\mathbb{E}\Bigg[\sup_{s\in[0,t]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\Bigg]\\
&\leq 8L^2(\delta+M+4)\int_{\delta}^{t}\mathbb{E}\Bigg[\sup_{w\in[0,r]}|X(w\wedge\tau_M)-\tilde{X}(w\wedge\tau_M)|^2\Bigg]dr\\
&\qquad+8L^2(\delta+M+4)\int_{\delta}^{t}\big(\mathbb{E}[|X(r)-\tilde{X}(r)|\mathbb{I}_{t<\tau_M}]+2\mathbb{P}(t\geq\tau_M)\big)^2dr\\
&\leq 8L^2(\delta+M+4)\int_{\delta}^{t}\mathbb{E}\Bigg[\sup_{w\in[0,r]}|X(w\wedge\tau_M)-\tilde{X}(w\wedge\tau_M)|^2\Bigg]dr\\
&\qquad+8L^2(\delta+M+4)\int_{\delta}^{t}2\mathbb{E}[|X(r)-\tilde{X}(r)|^2\mathbb{I}_{t<\tau_M}]+8\mathbb{P}(t\geq\tau_M)^2dr\\
&\leq 24L^2(\delta+M+4)\int_{\delta}^{t}\mathbb{E}\Bigg[\sup_{w\in[0,r]}|X(w\wedge\tau_M)-\tilde{X}(w\wedge\tau_M)|^2\Bigg]dr\\
&\qquad+64L^2(\delta+M+4)\int_{\delta}^{t}\mathbb{P}(t\geq\tau_M)^2dr\\
&\leq 24L^2(\delta+M+4)\int_{0}^{t}\mathbb{E}\Bigg[\sup_{w\in[0,r]}|X(w\wedge\tau_M)-\tilde{X}(w\wedge\tau_M)|^2\Bigg]dr\\
&\qquad+64L^2(\delta+M+4)\int_{\delta}^{t}\mathbb{P}\Big(\int_0^t|u(v)|^2dv\geq M\Big)^2dr\\
&\leq 24L^2(\delta+M+4)\int_{0}^{t}\mathbb{E}\Bigg[\sup_{w\in[0,r]}|X(w\wedge\tau_M)-\tilde{X}(w\wedge\tau_M)|^2\Bigg]dr\\
&\qquad+64L^2(\delta+M+4)\delta\mathbb{P}\Big(\int_0^T|u(v)|^2dv\geq M\Big)^2.
\end{aligned}$$

Exploiting the assumptions on the finiteness of the moments of exponential order of u together with the fact that
$\mathbb{P}\Big[\int_0^T|u(v)|^2dv\geq M\Big]=\mathbb{P}\Big[e^{c\int_0^T|u(v)|^2dv}\geq e^{cM}\Big]$,

using Markov inequality, we can rewrite the previous as:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[0,t]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\right]\\
&\le 24L^2(\delta+M+4)\int_0^t\mathbb{E}\left[\sup_{w\in[0,r]}|X(w\wedge\tau_M)-\tilde{X}(w\wedge\tau_M)|^2\right]dr\\
&\qquad+64L^2(\delta+M+4)\delta\frac{\mathbb{E}\left[e^{c\int_0^T|u(v)|^2dv}\right]^2}{e^{2cM}}\\
&\le 3B_M\int_0^t\mathbb{E}\left[\sup_{w\in[0,r]}|X(w\wedge\tau_M)-\tilde{X}(w\wedge\tau_M)|^2\right]dr+8B_M\delta\frac{V^2}{e^{2cM}},
\end{aligned}$$

where $B_M=8L^2(\delta+M+4)$ as above and $V=\mathbb{E}\left[e^{c\int_0^T|u(v)|^2dv}\right]$.
Now applying to the previous Gronwall Lemma, we get the estimation:

$$\mathbb{E}\left[\sup_{s\in[0,t]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\right]\le 8B_M\delta\frac{V^2}{e^{2cM}}e^{3B_Mt}.$$

Hence, for every $t\in[0,2\delta]$ we have the following

$$\mathbb{E}\left[\sup_{s\in[0,t]}|X(s\wedge\tau_M)-\tilde{X}(s\wedge\tau_M)|^2\right]\le 8B_M\delta V^2e^{-2cM+3B_M\delta}.$$

That, for $c>12\delta L^2$, thanks to monotone convergence theorem, at the limit for $M\to\infty$ gives:

$$\mathbb{E}\left[\sup_{s\in[0,2\delta]}|X(s)-\tilde{X}(s)|^2\right]\le 0,$$

and so $\mathbb{P}(X(t)=\tilde{X}(t),0\le t\le 2\delta)$.
Reasoning by induction repeating iteratively what we have done above, exploiting each time what we have found in the previous passage, we are able to show:
$\mathbb{P}(X(t)=\tilde{X}(t),0\le t\le T)=1$. □

Remark 2.6. *We can notice that the constant $c>0$, appearing in the constraint on the control, namely*

$$\mathbb{E}\left[e^{c\int_0^T|u(t)|^2dt}\right]<\infty,$$

is arbitrary. In fact, we can opportunely modify the proof by a suitable choice of $\delta>0$.

2.2 Delayed volatility coefficients

In order to prove Uniqueness of solutions, we can alternatively operate strengthening the hypothesis on the volatility coefficients .
As in the previous case, suppose to be interested in studying the non-linear SDE:

$$\begin{aligned} dX_t =& b(t, X, \mathrm{Law}(X(t)))dt + \sigma(t, X, \mathrm{Law}(X(t)))u(t)dt \\ &+ \sigma(t, X, \mathrm{Law}(X(t)))dW_t, \end{aligned} \tag{2.4}$$

where b, σ are predictable functionals on $[0, T] \times \mathcal{C}([0, T], \mathbb{R}^d) \times \mathcal{P}(\mathbb{R}^d)$ with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d_1}$ respectively, $X(0) = X_0$ fixed, such that $\mathbb{E}[|X_0|^2] < \infty$, and u a $\mathbb{R}^{d_1}$-valued $(\mathcal{F}_t)$-progressively measurable process such that

$$\mathbb{E}\left[\int_0^T |u(t)|^2 dt\right] = Q < \infty.$$

As previously, we consider the following Lipschitz and growth conditions on b and σ :

(L) There exists $L > 0$ such that for all $t \in [0, T]$, all $\phi, \psi \in \mathcal{C}([0, T], \mathbb{R}^d)$ all $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$

$$|b(t, \phi, \mu) - b(t, \psi, \nu)| + |\sigma(t, \phi, \mu) - \sigma(t, \psi, \nu)| \leq L\Big(\sup_{s\in[0,t]} |\phi(s) - \psi(s)| + d_{bL}(\mu, \nu)\Big).$$

(G) There exists a constant $K > 0$ such that for all $t \in [0, T]$, all $\phi \in \mathcal{C}([0, T], \mathbb{R}^d)$, all $\mu \in \mathcal{P}(\mathbb{R}^d)$

$$|b(t, \phi, \mu)| \leq K\Big(1 + \sup_{s\in[0,t]} |\phi(s)|\Big), \qquad |\sigma(t, \phi, \mu)| \leq K.$$

In addition, we suppose a condition of "delay" on σ:

(A) There exists $\delta > 0$, fixed, such that, for all $t \in [0, T]$, all $\phi, \psi \in \mathcal{C}([0, T], \mathbb{R}^d)$, all $\mu \in \mathcal{P}(\mathbb{R}^d)$

$$\sigma(t, \phi, \mu) = \sigma(t, \psi, \mu), \qquad \text{once } \phi(s) = \psi(s) \quad \forall s \in [0, t - \delta].$$

Let us notice that these conditions are sufficient to guarantee the finiteness of the moments of second order of the solutions of the SDE.†

†See Appendix B, in particular Remark B.3.

Remark 2.7. *We would like to underline that if $t \in [0,\delta)$ we will assume*

$$\sigma(t,\phi,\mu) = \sigma(t,\psi,\mu) \quad \textit{iff} \quad \phi(0) = \psi(0).$$

Proposition 2.8. *Grant conditions (L),(G) and(A). Let $((\Omega,\mathcal{F},\mathbb{P}),(\mathcal{F}_t)_{t\in[0,T]})$ be a stochastic basis satisfying the usual hypothesis and carrying a d_1-dimensional $(\mathcal{F})_t$-Wiener process $(W_t)_{t\in[0,T]}$ and let u be a $\mathbb{R}^{d_1}$-valued $(\mathcal{F}_t)$-progressively measurable process such that*

$$\mathbb{E}\left[\int_0^T |u(t)|^2 dt\right] < \infty.$$

Suppose that $X, \tilde{X}$ are solutions of (2.4), over the time interval $[0,T]$, under the control u, with initial condition $X(0) = X_0 = \tilde{X}(0)$ $\mathbb{P}$-almost surely. Then $X, \tilde{X}$ are indistinguishable, that is

$$\mathbb{P}(X(s) = \tilde{X}(s), \quad \forall s \in [0,T]) = 1.$$

Proof. The proof works in an inductive way. We exploit Property (A) on a sequence of intervals of amplitude δ to show that on each of them we have Uniqueness of solutions.
Set $\theta(t) = \mathrm{Law}(X(t)), \tilde{\theta}(t) = \mathrm{Law}(\tilde{X}(t)), t \in [0,T]$.
Let's start by considering the first time interval $[0,\delta]$. Suppose that $0 \le t < \delta$. Using Hölder's inequality, Doob's maximal inequality and Itô isometry, we obtain for all $t \in [0,\delta]$,

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]} |X(s)-\tilde{X}(s)|^2\right] &\le 4T\mathbb{E}\left[\int_0^t |b(r,X,\theta(r)) - b(r,\tilde{X},\tilde{\theta}(r))|^2 dr\right] \\ &+ 4\mathbb{E}\left[\int_0^t |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr \cdot \int_0^t |u(r)|^2 dr\right] \\ &+ 16\mathbb{E}\left[\int_0^t |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr\right].\end{aligned}$$

Thanks to hypothesis (L), applied to the first term on the right, and triangle inequality, we deduce:

$$\begin{aligned}&\le 8TL^2\mathbb{E}\left[\int_0^t \Big(\sup_{s\in[0,r]} |X(s)-\tilde{X}(s)|^2 + d_{bL}(\theta(r),\tilde{\theta}(r))^2\Big) dr\right] \\ &\quad + 4\mathbb{E}\left[\int_0^t (|\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\theta(r))| + |\sigma(r,\tilde{X},\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|)^2 dr \cdot \int_0^t |u(r)|^2 dr\right] \\ &\quad + 16\mathbb{E}\left[\int_0^t (|\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\theta(r))| + |\sigma(r,\tilde{X},\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|)^2 dr\right] = \diamondsuit.\end{aligned}$$

Exploiting hypothesis (A), in particular with regard to the Remark 2.7, and the fact that by hypothesis $X(0) = X_0 = \tilde{X}(0), \mathbb{P}$-a.s., we have that:

$$|\sigma(r, X, \theta(r)) - \sigma(r, \tilde{X}, \theta(r))| = 0, \quad \forall r \in [0, t], \quad \forall t \in [0, \delta], \quad \mathbb{P} - a.s.,$$

and so we are allowed to write:

$$\begin{aligned}
\diamondsuit &= 8TL^2\mathbb{E}\Bigg[\int_0^t \Big(\sup_{s\in[0,r]} |X(s) - \tilde{X}(s)|^2 + d_{bL}(\theta(r), \tilde{\theta}(r))^2\Big) dr\Bigg] \\
&\quad + 4\mathbb{E}\Bigg[\int_0^t d_{bL}(\theta(r), \tilde{\theta}(r))^2 dr \int_0^t |u(r)|^2 dr\Bigg] + 16\mathbb{E}\Bigg[\int_0^t d_{bL}(\theta(r), \tilde{\theta}(r))^2 dr\Bigg] \\
&= 8TL^2\mathbb{E}\Bigg[\int_0^t \sup_{s\in[0,r]} |X(s) - \tilde{X}(s)|^2 dr\Bigg] + 8TL^2 \int_0^t d_{bL}(\theta(r), \tilde{\theta}(r))^2 dr \\
&\quad + 4\mathbb{E}\Bigg[\int_0^t |u(r)|^2 dr\Bigg] \int_0^t d_{bL}(\theta(r), \tilde{\theta}(r))^2 dr + 16 \int_0^t d_{bL}(\theta(r), \tilde{\theta}(r))^2 dr \\
&= 8TL^2\mathbb{E}\Bigg[\int_0^t \sup_{s\in[0,r]} |X(s) - \tilde{X}(s)|^2 dr\Bigg] + (8TL^2 + 4Q + 16) \int_0^t d_{bL}(\theta(r), \tilde{\theta}(r))^2 dr,
\end{aligned}$$

with the notation $\mathbb{E}\Bigg[\int_0^\delta |u(r)|^2 dr\Bigg] \leq \mathbb{E}\Bigg[\int_0^T |u(r)|^2 dr\Bigg] = Q < \infty$.

Now, thanks to Property (2.1) of bounded Lipschitz metric and Cauchy-Schwarz inequality, we have:

$$d_{bL}(\theta(r), \tilde{\theta}(r))^2 \leq \mathbb{E}\Bigg[|X_r - \tilde{X}_r|\Bigg]^2 \leq \mathbb{E}\Bigg[|X_r - \tilde{X}_r|^2\Bigg],$$

and, naturally, we have:

$$\mathbb{E}\Bigg[|X_r - \tilde{X}_r|^2\Bigg] \leq \mathbb{E}\Bigg[\sup_{s\in[0,r]} |X_s - \tilde{X}_s|^2\Bigg].$$

Hence, we deduce that:

$$\mathbb{E}\Bigg[\sup_{s\in[0,t]} |X(s) - \tilde{X}(s)|^2\Bigg] \leq (16 + 16TL^2 + 4Q) \int_0^t \mathbb{E}\Bigg[\sup_{s\in[0,r]} |X(s) - \tilde{X}(s)|^2\Bigg] dr.$$

Applying Gronwall's Lemma, we finally get:

$$\forall t \in [0, \delta], \quad \mathbb{E}\Bigg[\sup_{s\in[0,t]} |X(s) - \tilde{X}(s)|^2\Bigg] \leq 0,$$

that is $\mathbb{P}(X(s) = \tilde{X}(s), \quad \forall s \in [0,\delta]) = 1$.
Now, we would like to extend the reasoning above to the time interval $[\delta, 2\delta]$.
We exploit the fact that $X(s) = \tilde{X}(s), \forall s \in [0,\delta] \quad \mathbb{P}$-a.s..
Let $t \in [\delta, 2\delta]$. We use in sequence Hölder's inequality, Doob's maximal inequality and Itô isometry, in order to obtain:

$$\mathbb{E}\left[\sup_{s\in[0,t]} |X(s)-\tilde{X}(s)|^2\right] \le 4T\mathbb{E}\left[\int_0^t |b(r,X,\theta(r)) - b(r,\tilde{X},\tilde{\theta}(r))|^2 dr\right]$$
$$+ 4\mathbb{E}\left[\int_0^t |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr \cdot \int_0^t |u(r)|^2 dr\right]$$
$$+ 16\mathbb{E}\left[\int_0^t |\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 dr\right] = \triangle.$$

Applying the hypothesis (L) to the first term and triangle inequality to the second and third

$$\triangle \le 8TL^2\mathbb{E}\left[\int_0^t \Big(\sup_{s\in[0,r]} |X(s)-\tilde{X}(s)|^2 + d_{bL}(\theta(r),\tilde{\theta}(r))^2\Big) dr\right]$$
$$+ 4\mathbb{E}\left[\int_0^t (|\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\theta(r))| + |\sigma(r,\tilde{X},\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|)^2 dr \cdot \int_0^t |u(r)|^2 dr\right]$$
$$+ 16\mathbb{E}\left[\int_0^t (|\sigma(r,X,\theta(r)) - \sigma(r,\tilde{X},\theta(r))| + |\sigma(r,\tilde{X},\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|)^2 dr\right] = \diamondsuit.$$

Since $\delta < t < 2\delta$ and remembering the previous step[‡], applying hypothesis (A), we can deduce: $\sigma(r,X,\theta(r)) = \sigma(r,\tilde{X},\theta(r)), \quad \forall r \in [0,t]$.
Furthermore, applying (L) we have: $|\sigma(r,\tilde{X},\theta(r)) - \sigma(r,\tilde{X},\tilde{\theta}(r))|^2 \le d_{bL}(\theta(r),\tilde{\theta}(r))^2$, $\forall r \in [0,t]$. We are able to write:

$$\diamondsuit \le 8TL^2\mathbb{E}\left[\int_0^t \sup_{s\in[0,r]} |X(s)-\tilde{X}(s)|^2 + d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr\right]$$
$$+ 4\mathbb{E}\left[\int_0^t |u(r)|^2 dr\right]\int_0^t d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr + 16\int_0^t d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr$$
$$= 8TL^2 \int_0^t \mathbb{E}\left[\sup_{s\in[0,r]} |X(s)-\tilde{X}(s)|^2\right] dr + (8TL^2 + 4Q + 16)\int_0^t d_{bL}(\theta(r),\tilde{\theta}(r))^2 dr$$
$$= (16TL^2 + 4Q + 16)\int_0^t \mathbb{E}\left[\sup_{s\in[0,r]} |X(s)-\tilde{X}(s)|^2\right] dr,$$

[‡] $\tilde{X}(s) = X(s), \forall s \in [0,\delta], \quad \mathbb{P} - a.s.$

where in the last passage we have exploited the fact, previously shown, that $d_{bL}(\theta(r),\tilde{\theta}(r))^2 \leq \mathbb{E}\left[\sup_{s\in[0,r]}|X(s)-\tilde{X}(s)|^2\right]$.
Applying Gronwall's Lemma to: $\forall t \in [0,2\delta]$

$$\mathbb{E}\left[\sup_{s\in[0,t]}|X(s)-\tilde{X}(s)|^2\right] \leq (16TL^2+4Q+16)\int_0^t \mathbb{E}\left[\sup_{s\in[0,r]}|X(s)-\tilde{X}(s)|^2\right]dr.$$

We get:

$$\mathbb{E}\left[\sup_{s\in[0,t]}|X(s)-\tilde{X}(s)|^2\right] \leq 0, \quad 0 \leq t \leq 2\delta,$$

that is:

$$\mathbb{P}(X(s)=\tilde{X}(s), \quad \forall s \in [0,2\delta]) = 1.$$

Reasoning iteratively on the various time interval of the type $[k\delta,(k+1)\delta]$, untill we cover the whole time interval $[0,T]$, we are able to show that:
$\forall k \in \{0,\ldots,l-1\}$

$$\mathbb{P}(X(s)=\tilde{X}(s), \quad \forall s \in [k\delta,(k+1)\delta]) = 1,$$

that is $X(s)=\tilde{X}(s) \quad \forall s \in [0,T] \quad \mathbb{P}$-a.s., once $\bigcup_{k=0}^{l-1}[k\delta,(k+1)\delta] \supset [0,T]$.

□

Appendix A

McKean Example of Interacting Diffusions

In this chapter we present an example which justifies the introduction of MV-SDEs.
Let's use a probabilistic method to study McKean's example of interacting diffusions. For this chapter we refer to [14].
Let $b : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded Lipschitz functional. Let u be a probability on $\mathbb{R}^d$ and denote with B the standard $\mathbb{R}^d$-Wiener measure. Construct on $(\mathbb{R}^d \times \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^d))^{N^*}$, with product measure $(u \otimes W)^{\otimes N^*}$, the processes $X^{i,N}, i = 1, \ldots, N$, satisfying:

$$\begin{aligned} dX_t^{i,N} &= dw_t^i + \frac{1}{N} \sum_{j=1}^{N} b(X_t^{i,N}, X_t^{j,N}) dt, \quad i = 1, \ldots, N; \\ X^{i,N}(0) &= x_0^i, \end{aligned} \tag{A.1}$$

where $x_0^i, (w^i), i \geq 1$, are the canonical coordinates on the product space $(\mathbb{R}^d \times \mathcal{C}_0)^{N^*}$. We will show that, when N goes to infinity, each $X^{i,N}$, has a natural limit $\bar{X}^i$. Each $\bar{X}^i$ will be independent copy of a new object: *"the non-linear process"*.
Let's give a brief description of the non-linear process.
Consider a filtered probability space, $(\Omega, \mathcal{F}, \mathcal{F}_t, (W_t)_{t\geq 0}, X_0, \mathbb{P})$, endowed with an $\mathbb{R}^d$-valued Wiener process $(W_t)_{t\geq 0}$, and an u-distributed, $\mathcal{F}_0$-measurable $\mathbb{R}^d$-valued random variable X_0. We study the equation:

$$\begin{aligned} dX_t &= \int_{\mathbb{R}^d} b(X_t, y) u_t(dy) dt + dW_t, \\ X(0) &= X_0, \quad u_t(dy) \text{ is the law of } X_t. \end{aligned} \tag{A.2}$$

Theorem A.1. *There is existence and uniqueness, trajectorial and in law, for the solution of SDE (A.2).*

Remark A.2. *We can notice that the non-linear process has time marginals which satisfy in a weak sense the non-linear equation:*

$$\partial_t u = \frac{1}{2}\Delta u - \operatorname{div}\Big(\int b(\cdot,y)u_t(dy)u\Big).$$

Indeed, for $f \in \mathcal{C}_c^2(\mathbb{R}^d)$, applying Itô's formula, we have:

$$f(X_t)-f(X_0) = \int_0^t \nabla f(X_s)dW_s + \int_0^t \Big(\frac{1}{2}\Delta f(X_s) + \int_{\mathbb{R}^d} b(X_s,y)u_s(dy)\nabla f(X_s)\Big)ds.$$

Integrating this, we get:

$$\int_{\mathbb{R}^d} f(x)(u_t(x)-u(x))dx = 0+\int_{\mathbb{R}^d}\int_0^t \Big(\frac{1}{2}\Delta f(x)+\int_{\mathbb{R}^d}, b(x,y)u_s(dy)\nabla f(x)\Big)u_s(x)dsdx,$$

that can be rewritten as:

$$\int_{\mathbb{R}^d}\int_0^t f(x)\partial u_s(x)dsdx = \int_{\mathbb{R}^d}\int_0^t \Big(\frac{1}{2}\Delta f(x)+\int_{\mathbb{R}^d} b(x,y)u_s(dy)\nabla f(x)\Big)u_s(x)dsdx.$$

Now, exploiting Divergence Theorem and the facts:

- $\operatorname{div}(gX) = \nabla g X + g\operatorname{div}(X)$
- $\operatorname{div}(fg) = f\operatorname{div}(g) + g\operatorname{div}(f) + 2\nabla f\nabla g$

together with the fact that f has compact support, we obtain:

$$\int\int_0^t f(x)\partial u_s(x)dsdx = \int\int_0^t \frac{1}{2}\Delta u_s(x)f(x) - \operatorname{div}\Big(\int_{\mathbb{R}^d} b(x,y)u_s(dy)u_s(x)\Big)f(x)dsdx$$

That is clearly a weak version of the equation above.

Let us now turn to the proof of Theorem A.1 .

Proof. We introduce the *Kantorovitch-Rubinstein* or *Wasserstein metric* on the set $\mathcal{P}(\mathcal{C})$ of probability measures on $\mathcal{C} = \mathcal{C}([0,T],\mathbb{R}^d)$, defined by

$$\begin{aligned} D_T(m_1,m_2) = \inf\Big\{ &\int (\sup_{s\le T}|X_s(\omega_1)-X_s(\omega_2)|\wedge 1)dm(\omega_1,\omega_2), \\ &m\in\mathcal{P}(\mathcal{C}\times\mathcal{C}),\quad p_1\circ m = m_1,\quad p_2\circ m = m_2\Big\}, \end{aligned} \tag{A.3}$$

where $(X_s)_{s\in[0,T]}$ is simply the canonical process on $\mathcal{C}$.
$D_T(\cdot,\cdot)$ is a complete metric on $\mathcal{P}(\mathcal{C})$, which gives to $\mathcal{P}(\mathcal{C})$ the topology of weak convergence. The proof of this fact can be found in [5].

Let $T > 0$. Define Φ the map which associates to $m \in \mathcal{P}(\mathcal{C})$ the law of the solution of:

$$X_t = X_0 + W_t + \int_0^t \Big(\int_{\mathcal{C}} b(X_s, w_s) dm(w)\Big) ds, \quad t \leq T. \tag{A.4}$$

Observe that this law does not depend on the specific choice of space Ω, we use.
If $m \in \mathcal{P}(\mathcal{C})$ is a fixed point of Φ, (A.4) defines a solution of (A.2), up to time T, and conversely, if $X_t, \quad t \leq T$, is a solution of (A.2), then its law on $\mathcal{C}([0,T], \mathbb{R})$, is a fixed point of Φ. There is a correspondence between our problem and a fixed point problem for Φ. We can exploit the following contraction lemma:

Lemma A.3. *For $t \leq T$,*

$$D_T(\Phi(m_1), \Phi(m_2)) \leq c_T \int_0^T D_u(m_1, m_2) du, \quad m_1, m_2 \in \mathcal{P}(\mathcal{C}),$$

where c_T is a constant and, $D_u(m_1, m_2) \quad (\leq D_T(m_1, m_2))$ is the distance between the images of m_1, and m_2 on $\mathcal{C}([0,u], \mathbb{R}^d)$.

Proof. Define the processes X^1 and X^2, as follows.

$$X^1_t = X_0 + W_t + \int_0^t \Big(\int_{\mathcal{C}} b(X^1_s, w_s) dm_1(w)\Big) ds, \quad t \leq T,$$
$$X^2_t = X_0 + W_t + \int_0^t \Big(\int_{\mathcal{C}} b(X^2_s, w_s) dm_2(w)\Big) ds, \quad t \leq T.$$

We have that

$$\begin{aligned}
&\sup_{s \leq t} |X^1_s - X^2_s| \\
&\leq \sup_{s \leq t} \Big| X_0 + W_s + \int_0^s \Big(\int_{\mathcal{C}} b(X^1_u, w_u) dm_1(w)\Big) du - \Big[X_0 + W_s \\
&\qquad + \int_0^s \Big(\int_{\mathcal{C}} b(X^2_u, w_u) dm_2(w)\Big) du\Big]\Big| \\
&\leq \sup_{s \leq t} \int_0^s \Big| \int_{\mathcal{C}} b(X^1_u, w_u) dm_1(w) - \int_{\mathcal{C}} b(X^2_u, w_u) dm_2(w)\Big| du \\
&\leq \int_0^t \Big| \int_{\mathcal{C}} b(X^1_u, w_u) dm_1(w) - \int_{\mathcal{C}} b(X^2_u, w_u) dm_2(w)\Big| du.
\end{aligned}$$

Since $b(\cdot, \cdot)$ is bounded and Lipschitz, exploiting triangle inequality, we can write

$$
\begin{aligned}
&\Big|\int_{\mathcal{C}} b(x,w_u)dm_1(w) - \int_{\mathcal{C}} b(y,w_u)dm_2(w)\Big| \\
&\leq \Big|\int_{\mathcal{C}} b(x,w_u)dm_1(w) \pm \int_{\mathcal{C}} b(y,w_u)dm_1(w) - \int_{\mathcal{C}} b(y,w_u)dm_2(w)\Big| \\
&\leq \Big|\int_{\mathcal{C}} b(x,w_u) - b(y,w_u)dm_1(w)| + |\int_{\mathcal{C}} b(y,w_u)d(m_1-m_2)(w)\Big| \\
&\leq \int_{\mathcal{C}} \Big|b(x,w_u) - b(y,w_u)|dm_1(w) + \int |b(y,w_u^1) - b(y,w_u^2)|dm(w_1,w_2) \\
&\leq K(|x-y| \wedge 1) + K\int |w_u^1 - w_u^2| \wedge 1 dm(w^1,w^2),
\end{aligned}
$$

where m is any coupling of m_1, m_2 on $\mathcal{C}([0,u],\mathbb{R}^d)$.
From this fact, we get

$$
\sup_{s\leq t} |X_s^1 - X_s^2| \leq K\int_0^t |X_s^1(\omega) - X_s^2(\omega)| \wedge 1 + K\int_0^t D_s(m_1,m_2)ds.
$$

Using Gronwall's Lemma, we have

$$
\sup_{s\leq t} |X_s^1 - X_s^2| \wedge 1 \leq Ke^{KT}\int_0^t D_s(m_1,m_2)ds,
$$

from which we can deduce

$$
\begin{aligned}
D_T(\Phi(m_1),\Phi(m_2)) = \inf\{&\int \Big(\sup_{s\leq t}|X_s^1 - X_s^2| \wedge 1\Big)dm(w_1,w_2) \\
&m \in \mathcal{P}(\mathcal{C}\times\mathcal{C}), \quad p_1 \circ m = \Phi(m_1), \quad p_2 \circ m = \Phi(m_2)\} \\
\leq \inf\{&\int Ke^{KT}\int_0^t D_s(m_1,m_2)dsdm(w_1,w_2) \\
&m \in \mathcal{P}(\mathcal{C}\times\mathcal{C}), \quad p_1 \circ m = \Phi(m_1), \quad p_2 \circ m = \Phi(m_2)\} \\
\leq Ke^{KT}&\int_0^t D_s(m_1,m_2)ds,
\end{aligned}
$$

and so the Lemma follows. □

From Lemma A.3, we can immediately deduce weak and strong uniqueness for the solutions of (A.2).
The existence part can be proved exploiting a standard contraction argument. Namely, for $T > 0$, and $m \in \mathcal{P}(\mathcal{C})$, iterating the lemma, we get:

$$
D_T(\Phi^{k+1}(m),\Phi^k(m)) \leq c_T\frac{T^k}{k!}D_T(\Phi(m),m)
$$

Hence, $(\Phi^k(m))_{k\in\mathbb{N}}$ is a Cauchy sequence, and converges to a fixed point of Φ: P_T. Now, if $T' < T$, the image of P_T on $\mathcal{C}([0,T'],\mathbb{R}^d)$ is still a fixed point, so the P_T are a consistent family, yielding a P on $\mathcal{C}([0,\infty],\mathbb{R}^d)$. This provides the required solution. □

Using Theorem A.1, we now introduce on $(\mathbb{R}^d \times \mathcal{C}_0)^{N^*}$, where we have constructed in (A.1) our interacting diffusions $X^{i,N}, \quad i = 1,\dots,N$, the processes $\bar{X}^i, i \geq 1$, solution of:

$$\begin{aligned} &\bar{X}^i_t = x^i_0 + w^i_t + \int_0^t \int_{\mathbb{R}^d} b(\bar{X}^i_s, y) u_s(dy) ds, \\ &u_s(dy) = \text{Law}(\bar{X}^i_s) \end{aligned} \tag{A.5}$$

Theorem A.4. *For any* $i \geq 1, \quad T > 0:$

$$\sup_N \sqrt{N}\mathbb{E}[\sup_{t\leq T} |X^{i,N}_t - \bar{X}^i_t|] < \infty. \tag{A.6}$$

Proof. We drop the superscript N for notational simplicity. We have:

$$\begin{aligned} X^i_t - \bar{X}^i{}_t &= \int_0^t \Big(\frac{1}{N}\sum_{j=1}^N b(X^i_s, X^j_s) - \int_{\mathbb{R}^d} b(\bar{X}^i_s, y) u_s(dy)\Big) ds \\ &= \int_0^t \frac{1}{N}\sum_{j=1}^N \{(b(X^i_s, X^j_s) - b(\bar{X}^i_s, X^j_s) + b(\bar{X}^i_s, X^j_s) - b(\bar{X}^i_s, \bar{X}^j_s)) \\ &\quad + (b(\bar{X}^i_s, \bar{X}^j_s) - \int_{\mathbb{R}^d} b(\bar{X}^i_s, y) u_s(dy))\} ds. \end{aligned}$$

Introducing the notation

$$b_s(x, x') = b(x, x') - \int_{\mathbb{R}^d} b(x, y) u_s(dy),$$

we see that:

$$\mathbb{E}[\sup_{0\leq t\leq T} |X^i_t - \bar{X}^i_t|] \leq K \int_0^T \Big(\mathbb{E}[|X^i_s - \bar{X}^i_s|] + \frac{1}{N}\sum_{j=1}^N \mathbb{E}[|X^j_s - \bar{X}^j_s|] + \mathbb{E}[|\frac{1}{N}\sum_{j=1}^N b_s(\bar{X}^i_s, \bar{X}^j_s)|]\Big) ds.$$

Summing the previous inequality over $i = 1,\dots,N$ and using symmetry, we find:

$$\begin{aligned} N\mathbb{E}[\sup_{0\leq t\leq T} |X^i_t - \bar{X}^i_t|] &= \sum_{i=1}^N \mathbb{E}[\sup_{0\leq t\leq T} |X^i_t - \bar{X}^i{}_t|] \\ &\leq K' \int_0^T \sum_{i=1}^N \Big(\mathbb{E}[|X^i_s - \bar{X}^i_s|] + \mathbb{E}[|\frac{1}{N}\sum_{j=1}^N b_s(\bar{X}^i_s, \bar{X}^j_s)|]\Big) ds. \end{aligned}$$

Applying Gronwall's lemma, and symmetry, we find:

$$\mathbb{E}[\sup_{0\le t\le T}|X_t^i-\bar{X}_t^i|]\le K(T)\int_0^T\Big(\mathbb{E}[|\frac{1}{N}\sum_{j=1}^N b_s(\bar{X}_s^i,\bar{X}_s^j)|]\Big)ds.$$

Our claim will follow provided that we can show:

$$\mathbb{E}\Big[|\frac{1}{N}\sum_{j=1}^N b_s(\bar{X}_s^i,\bar{X}_s^j)|\Big]\le\frac{C(T)}{\sqrt{N}}.$$

We can notice that

$$\mathbb{E}\Big[\big(\frac{1}{N}\sum_{j=1}^N b_s(\bar{X}_s^i,\bar{X}_s^j)\big)^2\Big]=\frac{1}{N^2}\mathbb{E}\Big[\sum_{j,k=1}^N b_s(\bar{X}_s^i,\bar{X}_s^j)b_s(\bar{X}_s^i,\bar{X}_s^k)\Big]$$

and, for $j\neq k$, we have:

$$\begin{aligned}&\mathbb{E}\Big[b_s(\bar{X}_s^i,\bar{X}_s^j)b_s(\bar{X}_s^i,\bar{X}_s^k)\Big]\\&=\mathbb{E}\Big[\Big(b(\bar{X}_s^i,\bar{X}_s^j)-\int_{\mathbb{R}^d}b(\bar{X}_s^i,y)u_s(dy)\Big)\Big(b(\bar{X}_s^i,\bar{X}_s^k)-\int_{\mathbb{R}^d}b(\bar{X}_s^i,y)u_s(dy)\Big)\Big]\\&=\mathbb{E}\Big[b(\bar{X}_s^i,\bar{X}_s^j)b(\bar{X}_s^i,\bar{X}_s^k)-\int_{\mathbb{R}^d}b(\bar{X}_s^i,y)u_s(dy)(b(\bar{X}_s^i,\bar{X}_s^j)+b(\bar{X}_s^i,\bar{X}_s^k))\\&\qquad+(\int_{\mathbb{R}^d}b(\bar{X}_s^i,y)u_s(dy))^2\Big]=0,\end{aligned}$$

and so we get:

$$\mathbb{E}\Big[\big(\frac{1}{N}\sum_{j=1}^N b_s(\bar{X}_s^i,\bar{X}_s^j)\big)^2\Big]=\frac{1}{N^2}\mathbb{E}\Big[\sum_{j=1}^N b_s(\bar{X}_s^i,\bar{X}_s^j)^2\Big]\le\frac{1}{N}\mathbb{E}\Big[b_s(\bar{X}_s^i,\bar{X}_s^j)^2\Big]=\frac{\chi^2}{N}.$$

Hence, we can deduce that:

$$\mathbb{E}\Big[|\frac{1}{N}\sum_{j=1}^N b_s(\bar{X}_s^i,\bar{X}_s^j)|\Big]\le\mathbb{E}\Big[\big(\frac{1}{N}\sum_{j=1}^N b_s(\bar{X}_s^i,\bar{X}_s^j)\big)^2\Big]^{\frac{1}{2}}\le\Big(\frac{\chi^2}{N}\Big)^{\frac{1}{2}}=\frac{\chi}{\sqrt{N}},$$

from which our claim follows. □

$\mathcal{P}(E)$ denotes, here, the set of probability measures on E, where E is a separable metric space.

Definition A.5. *Let E a separable metric space, u_N a sequence of symmetric probabilities on E^N. We say that u_N is $u-$**chaotic**, u probability on E, if for $\phi_1,\dots,\phi_k\in\mathcal{C}_b(E)$, $k\ge1$,*

$$\lim_{N\to\infty}\langle u_N,\phi_1\otimes\cdots\otimes\phi_k\otimes1\cdots\otimes1\rangle=\prod_{i=1}^k\langle u,\phi_i\rangle. \tag{A.7}$$

The notion of $u-$chaotic means that the empirical measures of the coordinate variables of E^N, under u_N tend to concentrate near u, as the next proposition shows. This is a type of law of large numbers.
Condition (A.7) can also be restated as the convergence of the projection of u_N as E^k to $u^{\otimes k}$ when N goes to infinity. In the coming proposition we suppose u_N symmetric.

Proposition A.6.

i) u_N *is* u-chaotic *is equivalent to* $\bar{X}_N = \frac{1}{N}\sum_1^N \delta_{X_i}$ *(* $\mathcal{P}(E)$*-valued random variables on* (E^N, u_N)*,* X_i *canonical coordinates on* E^N*) converges in law to the constant random variable* u*.*
It is also equivalent to condition (A.7), with $k = 2$*.*

ii) When E *is a Polish space, the* $\mathcal{P}(E)$*-valued variables* $\bar{X}_N$ *are tight if and only if the laws on* E *of* X_1 *under* u_N *are tight.*

Proof. Let's start by proving i).
First, suppose u_N satisfies (A.7) with $k = 2$, and consequently with $k = 1$ as well. Take $\phi \in \mathcal{C}_b(E)$, we want to compute $\mathbb{E}[\langle \bar{X}_N - u, \phi\rangle^2]$. By definition, we have:

$$\langle \bar{X}_N - u, \phi\rangle = \langle \frac{1}{N}\sum_{i=1}^N \delta_{X_i} - u, \phi\rangle = \frac{1}{N}\sum_{i=1}^N \phi(X_i) - \langle u, \phi\rangle$$

Hence, we get:

$$\begin{aligned}
&\mathbb{E}[\langle \bar{X}_N - u, \phi\rangle^2] = \mathbb{E}\Big[\Big(\frac{1}{N}\sum_{i=1}^N \phi(X_i) - \langle u, \phi\rangle\Big)^2\Big] \\
&= \mathbb{E}\Big[\frac{1}{N^2}\sum_{i,j=1}^N \phi(X_i)\phi(X_j) - 2\langle u, \phi\rangle \frac{1}{N}\sum_{i=1}^N \phi(X_i)\Big] + \langle u, \phi\rangle^2 \\
&= \frac{1}{N^2}\sum_{i,j=1}^N \mathbb{E}[\phi(X_i)\phi(X_j)] - 2\langle u, \phi\rangle \frac{1}{N}\sum_{i=1}^N \mathbb{E}[\phi(X_i)] + \langle u, \phi\rangle^2 = \triangle.
\end{aligned}$$

Now, exploiting the symmetry and the fact that $\{X_i\}_i$ are i.d., we have:

$$\begin{aligned}
\triangle &\leq \frac{1}{N^2}\sum_{i=1}^N \mathbb{E}[\phi(X_i)^2] + \frac{1}{N^2}\sum_{i,j=1,(i\neq j)}^N \mathbb{E}[\phi(X_i)\phi(X_j)] - 2\langle u, \phi\rangle \frac{1}{N}\sum_{i=1}^N \mathbb{E}[\phi(X_i)] + \langle u, \phi\rangle^2 \\
&= \frac{1}{N}\mathbb{E}[\phi(X_1)^2] + \frac{N-1}{N}\mathbb{E}[\phi(X_1)\phi(X_2)] - 2\langle u, \phi\rangle\mathbb{E}[\phi(X_1)] + \langle u, \phi\rangle^2.
\end{aligned}$$

Now, we can write:

$$\mathbb{E}[\phi(X_1)^2] = \langle u_N, \phi^2 \otimes 1 \cdots \otimes 1\rangle \longrightarrow \langle u, \phi^2\rangle \quad \text{by (A.7) with } k = 1,$$

and so we get

$$\frac{1}{N}\mathbb{E}[\phi(X_1)^2] \longrightarrow 0.$$

Moreover, we have:

$$\mathbb{E}[\phi(X_1)] = \langle u_N, \phi \otimes 1 \cdots \otimes 1\rangle \longrightarrow \langle u, \phi\rangle \quad \text{by (A.7) with } k=1,$$

and

$$\mathbb{E}[\phi(X_1)\phi(X_2)] = \langle u_N, \phi \otimes \phi \otimes 1 \cdots \otimes 1\rangle \longrightarrow \langle u, \phi\rangle^2 \quad \text{by (A.7) with } k=2.$$

This implies that $\bar{X}_N$ converges in law to the constant random variable equal to u.
Conversely, suppose $\bar{X}_N$ converges in law to the constant u. Exploiting triangle inequality, we can write:

$$\begin{aligned}
&|\langle u_N, \phi_1 \otimes \cdots \otimes \phi_k \otimes 1 \cdots \otimes 1\rangle - \prod_{i=1}^{k}\langle u, \phi_i\rangle| \\
&\leq |\langle u_N, \phi_1 \otimes \cdots \otimes \phi_k \otimes 1 \cdots \otimes 1\rangle - \langle u_N, \prod_{i=1}^{k}\langle \bar{X}_N, \phi_i\rangle\rangle| \\
&\quad + |\langle u_N, \prod_{i=1}^{k}\langle \bar{X}_N, \phi_i\rangle\rangle - \prod_{i=1}^{k}\langle u, \phi_i\rangle|
\end{aligned} \tag{A.8}$$

The second term of the previous inequality goes to zero, since by hypothesis $\bar{X}_N$ converges in law to the constant u.
The first term, using symmetry, can be rewritten as:

$$|\langle u_N, \frac{1}{N!}\sum_{\sigma \in S_N} \phi_1(X_{\sigma(1)}) \dots \phi_k(X_{\sigma(k)}) - \prod_{i=1}^{k}\langle \bar{X}_N, \phi_i\rangle\rangle|$$

Observe now that, if $M \geq \| \phi_i \|_\infty, 1 \leq i \leq k$, we have:

$$\begin{aligned}
&\sup_{E^N}\Big|\frac{1}{N!}\sum_{\sigma \in S_N} \phi_1(X_{\sigma(1)}) \dots \phi_k(X_{\sigma(k)}) - \prod_{i=1}^{k}\langle \bar{X}_N, \phi_i\rangle\Big| \\
&\leq M^k\Big[\Big(\frac{(N-k)!}{N!} - \frac{1}{N^k}\Big) \cdot \frac{N!}{(N-k)!} + \frac{1}{N^k}\Big(N^k - \frac{N!}{(N-k)!}\Big)\Big] \\
&= 2M^k\Big(1 - \frac{N!}{N^k(N-k)!}\Big) \longrightarrow 0
\end{aligned}$$

Here we simply used that there are $N!/(N-k)!$ injections from $\{1, \dots, k\}$ into $\{1, \dots, N\}$ each of them has weight $(N-k)!/N!$ in the first sum and $1/N^k$ in

the second sum, and in the second sum there are also $N^k - N!/(N-k)!$ terms where repetitions of coordinates occur. So we see that the first term of (A.8) goes to zero, and this proves i).
Let's prove ii) .

Definition A.7. *For a probability $\mathbb{Q}(dm)$ on $\mathcal{P}(E)$, define the **intensity**, $I(\mathbb{Q})$, as the probability measure:*

$$\langle I(\mathbb{Q}), f\rangle = \int_{\mathcal{P}(E)} \langle m, f\rangle d\mathbb{Q}(m) = \int_{\mathcal{P}(E)} \int_E f(x)dm(x)d\mathbb{Q}(m), \tag{A.9}$$

for $f \in B(E)$, the space of bounded functions on E.

We would like to show a more general fact.

Proposition A.8. *Tightness for a family of measure $\mathbb{Q}$ on $\mathcal{P}(E)$ is equivalent to the tightness of their intensity measures $I(\mathbb{Q})$ on E.*

Our claim ii) follows directly from the proposition above. In fact, in our situation of ii), by symmetry, the intensity measure of the law of $\bar{X}_N$ is just the law of X_1, under E_N.
Indeed, with $\mathbb{Q}_{\bar{X}_N}$ law of $\bar{X}_N$, we have :

$$\begin{aligned}
\langle I(\mathbb{Q}_{\bar{X}_N}), f\rangle &= \int_{\mathcal{P}(E)} \langle m, f\rangle d\mathbb{Q}_{\bar{X}_N}(m) = \int_{\mathcal{P}(E)} \langle m, f\rangle \frac{1}{N}\sum_{i=1}^N \delta_{X_i}(m) \\
&= \frac{1}{N}\sum_{i=1}^N \int_{\mathcal{P}(E)} \int_E f(x)dm(x)\delta_{X_i}(m) = \frac{1}{N}\sum_{i=1}^N \int_E f(x)dm_{X_1}(x) \\
&= \int_E f(x)dm_{X_1}(x) = \mathbb{E}[f(X_1)] = \langle \mathbb{P}\circ (X_1)^{-1}, f\rangle,
\end{aligned}$$

where $\mathbb{P}\circ (X_1)^{-1}$ stands for the law of the projection on the first coordinates of u_N.

Let's prove Proposition A.8.

Proof. The map that associates to a $\mathbb{Q} \in \mathcal{P}(\mathcal{P}(E))$ its intensity, $I(\mathbb{Q})$, is clearly continuous for the respective weak convergence topologies. So, the statement above will follow if we prove that whenever $(I_n)_{n\in\mathbb{N}}$, $I_n = I(\mathbb{Q}_n)$, is tight, then $(\mathbb{Q}_n)_{n\in\mathbb{N}}$ is tight.
For each $\epsilon > 0$, denote by K_ϵ a compact subset of E, with $I_n(K_\epsilon^c) \leq \epsilon$, for every $n \in \mathbb{N}$.
Now for $\epsilon, \eta > 0$, and any $n \in \mathbb{N}$, exploiting Markov inequality, we get:

$$\mathbb{Q}_n(\{m \in \mathcal{P}(E)/m(K_{\epsilon\eta}^c) \geq \eta\}) \leq \frac{1}{\eta} I_n(K_{\epsilon\eta}^c) \leq \epsilon.$$

It follows that :

$$\mathbb{Q}_n\left(\bigcup_{k\geq 1}\left\{m\in\mathcal{P}(E)/m(K^c_{\epsilon\frac{2^{-k}}{k}})\geq\frac{1}{k}\right\}\right)\leq\sum_{k\geq 1}^{\infty}\epsilon 2^{-k}\leq\epsilon.$$

This means that $\mathbb{Q}_n$ puts a mass greater or equal to $1-\epsilon$ on the compact subset of $\mathcal{P}(E)$, $\bigcap_{k\geq 1}\{m\in\mathcal{P}(E)/m(K^c_{\epsilon\frac{2^{-k}}{k}})\geq\frac{1}{k}\}$.
This proves the $(\mathbb{Q}_n)_{n\in\mathbb{N}}$ are tight. □

□

We are now interested in studying the relation between weak convergence of a sequence of measures on $\mathcal{P}(E)$, namely $(\mathbb{Q}_n)_{n\in\mathbb{N}}$, and the weak convergence of the sequence of their intensities: $(I_n)_n=(I(\mathbb{Q}_n))_n$.
Let $(\mathbb{Q}_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $\mathcal{P}(E)$.
We say that $(\mathbb{Q}_n)_{n\in\mathbb{N}}$ *converges weakly to* $\mathbb{Q}$,

$$\mathbb{Q}_n\longrightarrow^{weakly}_{n\to\infty}\mathbb{Q}\qquad\text{if}\quad\forall f\in\mathcal{C}_b(\mathcal{P}(E),\mathbb{R})$$
$$\langle\mathbb{Q}_n,f\rangle=\int_{\mathcal{P}(E)}f(m)d\mathbb{Q}_n(m)\longrightarrow_{n\to\infty}\langle\mathbb{Q},f\rangle=\int_{\mathcal{P}(E)}f(m)d\mathbb{Q}(m).$$

Considering the sequence of the intensities, $(I_n)_n=(I(\mathbb{Q}_n))_n$, which is a sequence of probability measure on E, we say that $(I(\mathbb{Q}_n))_{n\in\mathbb{N}}$ *converges weakly to* $I(\mathbb{Q})$,

$$I(\mathbb{Q}_n)\longrightarrow^{weakly}_{n\to\infty}I(\mathbb{Q})\qquad\text{if}\quad\forall g\in\mathcal{C}_b(E,\mathbb{R})$$
$$\langle I(\mathbb{Q}_n),g\rangle=\int_{\mathcal{P}(E)}\langle m,g\rangle d\mathbb{Q}_n(m)\longrightarrow_{n\to\infty}\langle I(\mathbb{Q}),g\rangle=\int_{\mathcal{P}(E)}\langle m,g\rangle d\mathbb{Q}(m).$$

It is clear that $\mathbb{Q}_n\longrightarrow^{weakly}_{n\to\infty}\mathbb{Q}$ implies that $I(\mathbb{Q}_n)\longrightarrow^{weakly}_{n\to\infty}I(\mathbb{Q})$.
In fact $\forall g\in\mathcal{C}_b(E,\mathbb{R})$ we have:

$$\begin{aligned}\langle I(\mathbb{Q}_n),g\rangle&=\int_{\mathcal{P}(E)}\langle m,g\rangle d\mathbb{Q}_n(m)=\int_{\mathcal{P}(E)}G(m)d\mathbb{Q}_n(m)\longrightarrow_{n\to\infty}\langle I(\mathbb{Q}),G\rangle\\&=\int_{\mathcal{P}(E)}\langle m,G\rangle d\mathbb{Q}(m),\end{aligned}$$

where we have exploited the fact that for $g\in\mathcal{C}_b(E,\mathbb{R})$, G defined by $G(m)=\langle m,g\rangle=\int_E g(x)m(dx)$ is in $\mathcal{C}_b(\mathcal{P}(E),\mathbb{R})$.
We want to check if it is true that:
$I(\mathbb{Q}_n)\longrightarrow^{weakly}_{n\to\infty}I(\mathbb{Q})$ implies $\mathbb{Q}_n\longrightarrow^{weakly}_{n\to\infty}\mathbb{Q}$.

Under the hypothesis that $(I(\mathbb{Q}_n))_{n\in\mathbb{N}}$ converges weakly to $I(\mathbb{Q})$, hence $\forall g \in \mathcal{C}_b(E,\mathbb{R})$

$$\langle I(\mathbb{Q}_n), g\rangle = \int_{\mathcal{P}(E)} \langle m, g\rangle d\mathbb{Q}_n(m) \longrightarrow_{n\to\infty} \langle I(\mathbb{Q}), g\rangle = \int_{\mathcal{P}(E)} \langle m, g\rangle d\mathbb{Q}(m)$$

we can deduce that $(I_n)_{n\in\mathbb{N}}$ is tight.
Exploiting Proposition A.8 we have that also the sequence $(\mathbb{Q}_n)_n$ is tight.
Now we consider an arbitrary sub-sequence of $(\mathbb{Q}_n)_n$, $(\mathbb{Q}_{n_k})_k$.
$(\mathbb{Q}_{n_k})_k$ is tight because of the tightness of $(\mathbb{Q}_n)_n$. By Prokhorov Theorem there exists a sub-sub-sequence $(\mathbb{Q}_{n_{k_l}})_l$ weakly converging to some probability $\mathbb{B}$. Because of the hypothesis: $(I(\mathbb{Q}_n))_{n\in\mathbb{N}}$ converges weakly to $I(\mathbb{Q})$, it must be necessarily $I(\mathbb{B}) = I(\mathbb{Q})$.
The problem lays in the fact that $I(\mathbb{B}) = I(\mathbb{Q})$, does not imply that $\mathbb{B} = \mathbb{Q}$.

We can check this in simpler case. Suppose $E = \{0, 1\}$.
In this case $\mathcal{P}(E) = \{m_\alpha = (m_\alpha(0), m_\alpha(1)) = (\alpha, 1-\alpha) : \alpha \in [0, 1]\}$.
Let $\mathbb{Q}$ be a probability measure on $\mathcal{P}(E)$, we define its intensity by:
$\forall f \in B(E)$

$$\begin{aligned}\langle I(\mathbb{Q}), f\rangle &= \int_{\mathcal{P}(E)} \alpha f(0) + (1-\alpha) f(1) \mathbb{Q}(dm_\alpha)\\ &= \int_{\mathcal{P}(E)} f(1) + \alpha(f(0) - f(1))\mathbb{Q}(dm_\alpha).\end{aligned}$$

Considering two probability measure on $\mathcal{P}(E)$, respectively $\mathbb{Q}_1 = \delta_{m_{\frac{1}{2}}}$ and $\mathbb{Q}_2 = \frac{1}{2}(\delta_{m_0} + \delta_{m_1})$, have that:

$$\begin{aligned}\langle I(\mathbb{Q}_1), f\rangle &= \int_{\mathcal{P}(E)} f(1) + \alpha(f(0) - f(1))\delta_{m_{\frac{1}{2}}}(dm_\alpha) = f(1) + \frac{1}{2}(f(0) - f(1))\\ &= \frac{1}{2}(f(0) + f(1)),\end{aligned}$$

$$\begin{aligned}\langle I(\mathbb{Q}_2), f\rangle &= \int_{\mathcal{P}(E)} f(1) + \alpha(f(0) - f(1))\frac{1}{2}(\delta_{m_0} + \delta_{m_1})(dm_\alpha)\\ &= \frac{1}{2}(f(1) + f(1) + (f(0) - f(1)))\\ &= \frac{1}{2}(f(0) + f(1)),\end{aligned}$$

that is $\forall f \in B(E) : \langle I(\mathbb{Q}_1), f\rangle = \langle I(\mathbb{Q}_2), f\rangle$ and so $I(\mathbb{Q}_1) = I(\mathbb{Q}_2)$.
Nevertheless, clearly, we have $\mathbb{Q}_1 \neq \mathbb{Q}_2$. In fact, if we consider
$g : \mathcal{P}(E) \longrightarrow \mathbb{R}, g \in \mathcal{C}_b(E), g \geq 0, g(m_0) = 1, supp(g) \subset \{m_\alpha : \alpha \in [0, 1/4]\}$,

$$\langle \mathbb{Q}_1, g \rangle = 0,$$
$$\langle \mathbb{Q}_2, g \rangle = \frac{1}{2},$$

that is $\langle \mathbb{Q}_1, g \rangle \neq \langle \mathbb{Q}_2, g \rangle$, and so $\mathbb{Q}_1 \neq \mathbb{Q}_2$.
The argument can be simply extended to any E finite and let us think that $I(\mathbb{Q})$ is a sort of "mean" of $\mathbb{Q}$.

Exploiting the example given above we can formulate a counterexample for the *Proposition*:
$I(\mathbb{Q}_n) \longrightarrow_{n\to\infty}^{weakly} I(\mathbb{Q})$ implies $\mathbb{Q}_n \longrightarrow_{n\to\infty}^{weakly} \mathbb{Q}$.
Consider the sequence of probability measures on $\mathcal{P}(E)$, $(\mathbb{Q}_n)_{n\in\mathbb{N}}$, given by:

$$\mathbb{Q}_{2k} = \delta_{m_{\frac{1}{2}}}$$
$$\mathbb{Q}_{2k+1} = \frac{1}{2}(\delta_{m_0} + \delta_{m_1}), \qquad k = 0, 1, 2, \dots$$

We have that clearly $(\mathbb{Q}_n)_{n\in\mathbb{N}}$ can not converge weakly to any $\mathbb{Q} \in \mathcal{P}(\mathcal{P}(E))$.
As we have seen above, it is true that $I(\mathbb{Q}_n) = I(\mathbb{Q}), \forall n \in \mathbb{N}$, and so
$I(\mathbb{Q}_n) \longrightarrow_{n\to\infty}^{weakly} I(\mathbb{Q})$, with $\mathbb{Q} = \delta_{m_{\frac{1}{2}}}$.
We have found the counterexample we were looking for:
A sequence in $\mathcal{P}(\mathcal{P}(E))$ for which we have: $I(\mathbb{Q}_n) \longrightarrow_{n\to\infty}^{weakly} I(\mathbb{Q})$, but not $\mathbb{Q}_n \longrightarrow_{n\to\infty}^{weakly} \mathbb{Q}$.

Appendix B

Some notes on SDEs

We want to recall some well-known results related to linear SDEs. We need to exploit them in order to show existence and uniqueness of solutions in our non-linear environment.

Let $T > 0$ be a time horizon and $d, d_1 \in \mathbb{N}$.
Let $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t\in[0,T]})$ be a stochastic basis satisfying the usual hypotheses and carrying a d_1-dimensional $(\mathcal{F})_t$-Wiener process $(W_t)_{t\in[0,T]}$.

Theorem B.1. *Consider the following SDE:*

$$dX_t = b(t, X, \theta_t)dt + \sigma(t, X, \theta_t)u(t)dt + \sigma(t, X, \theta_t)dW_t, \tag{B.1}$$

with the initial condition $X(0) = X_0$, such that $\mathbb{E}[|X_0|^2] < \infty$, $\theta \in \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$ fixed and u a $\mathbb{R}^{d_1}$-valued $(\mathcal{F}_t)$-progressively measurable process such that

$$\mathbb{E}\left[\int_0^T |u(t)|^2 dt\right] = Q < \infty.$$

Under hypotheses (L) *and* (G) *on b and σ, we have that Equation (B.1) has a unique strong solution, whose moment of second order satisfies the following:*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^2\Big] \le Re^{8K^2T^2} = R' < \infty, \tag{B.2}$$

with $R = 4\big(\mathbb{E}[|X_0|^2] + TK^2\mathbb{E}[\int_0^T |u_r|^2 dr] + 4TK^2 + 2K^2T^2\big)$.

Remark B.2. *We can notice that the constant R' does not depend on the specific $\theta \in \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$, that is involved in the equation.*

Proof. Let's start by proving the uniqueness of solutions. Let $X = (X_t)_{t\in[0,T]}$ and $\tilde{X} = (\tilde{X}_t)_{t\in[0,T]}$ be strong solutions of Equation (B.1) for a fixed $\theta \in \mathcal{P}_1(\mathbb{R}^d)$. For $M \in \mathbb{N}$, define an $(\mathcal{F}_t)$-stopping time τ_M by

$$\tau_M(\omega) := \inf\{t \in [0, T] : |X(t, \omega)| \wedge |\tilde{X}(t, \omega)| \wedge \int_0^t |u(s, \omega)|^2 ds \ge M\},$$

with $\inf \varnothing = \infty$. Observe that $\mathbb{P}(\tau_M \le T) \to 0$ as $M \to \infty$, since $X, \tilde{X}$ are continuous processes and $\mathbb{E}[\int_0^T |u(s)|^2 ds] < Q$.
Using Hölder's inequality, Doob's maximal inequality, the Itô isometry, and condition (L), we obtain for $M \in \mathbb{N}$, all $t \in [0,T]$:

$$\begin{aligned}
\mathbb{E}\Big[\sup_{s\in[0,t]} |X_{s\wedge\tau_M} - \tilde{X}_{s\wedge\tau_M}|^2\Big] &\le 4T\mathbb{E}\left[\int_0^{t\wedge\tau_M} |b(r,X,\theta_r) - b(r,\tilde{X},\theta_r)|^2 dr\right] \\
&\quad + 4\mathbb{E}\left[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta_r) - \sigma(r,\tilde{X},\theta_r)|^2 dr \int_0^{t\wedge\tau_M} |u(r)|^2 dr\right] \\
&\quad + 16\mathbb{E}\left[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta_r) - \sigma(r,\tilde{X},\theta_r)|^2 dr\right] \\
&\le 4T\mathbb{E}\left[\int_0^{t\wedge\tau_M} |b(r,X,\theta_r) - b(r,\tilde{X},\theta_r)|^2 dr\right] \\
&\quad + (4M+16)\mathbb{E}\left[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta_r) - \sigma(r,\tilde{X},\theta_r)|^2 dr\right] \\
&\le (4T+4M+16)\mathbb{E}\left[\int_0^{t\wedge\tau_M} \sup_{s\in[0,r]} |X_s - \tilde{X}_s|^2 dr\right] \\
&\le (4T+4M+16)\int_0^t \mathbb{E}\left[\sup_{s\in[0,r]} |X_{s\wedge\tau_M} - \tilde{X}_{s\wedge\tau_M}|^2\right] dr.
\end{aligned}$$

An application of Gronwall's lemma yields that

$$\mathbb{E}\Big[\sup_{s\in[0,T]} |X_{s\wedge\tau_M} - \tilde{X}_{s\wedge\tau_M}|^2\Big] = 0,$$

hence $\mathbb{P}(X(t) = \tilde{X}(t), \text{ for all } t < \tau_M) = 1$ for all $M \in \mathbb{N}$. This implies the assertion since $\tau_M \nearrow \infty$ as $M \to \infty$ $\mathbb{P}$-almost surely.

Now, we are interested in showing the existence of solutions for Equation (B.1) for a fixed $\theta \in \mathcal{P}_1(\mathbb{R}^d)$.
Denote with $M_2[0,T]$ the vector space of $\mathbb{R}^d$-valued, progressively measurable processes $Y = (Y_t)_{t\in[0,T]}$, such that: $\mathbb{E}[\sup_{t\in[0,T]} |X_t|^2] < \infty$. On this space we consider the following metric

$$\|X - Y\|_{M_2} := \sqrt{\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t - Y_t|^2\Big]}, \qquad X, Y \in M_2[0,T].$$

For $M \in \mathbb{N}$, define an $(\mathcal{F}_t)-$stopping time τ_M by

$$\begin{aligned} &\tau_0(\omega) := 0, \\ &\tau_M(\omega) := \inf\{t \in [0,T] : \int_0^t |u(s,\omega)|^2 ds \geq M\}, \qquad M \geq 1, \end{aligned}$$

with $\inf \varnothing = \infty$. Observe that τ_M is $\mathbb{P}-$a.s. increasing in M and that $\mathbb{P}(\tau_M \leq T) \to 0$ as $M \to \infty$, since $\mathbb{E}[\int_0^T |u(s)|^2 ds] = Q < \infty$.
For $M \in \mathbb{N}$, we define the control u_M as:

$$u_M(t,\omega) := \mathbb{I}_{[0,\tau_M(\omega)]}(t) u(t,\omega).$$

Hence, for all $M \in \mathbb{N}$, we can define a new SDE:

$$\begin{aligned} &dX_t = b(t,X,\theta_t)dt + \sigma(t,X,\theta_t)u_M(t)dt + \sigma(t,X,\theta_t)dW_t, \\ &X(0) = X_0. \end{aligned} \tag{B.3}$$

We are going to show that, for all $M \in \mathbb{N}$, the Equation (B.3) has a unique strong solution $X^M = (X_t^M)_{t\in[0,T]}$.
Let's start by proving the uniqueness part.
Consider $X^M = (X_t^M)_{t\in[0,T]}$ and $\tilde{X}^M = (\tilde{X}_t^M)_{t\in[0,T]}$ strong solutions of (B.3).
For $L \in \mathbb{N}$, define an $(\mathcal{F}_t)$-stopping time τ_L by

$$\tau_L(\omega) := \inf\{t \in [0,T] : |X^M(t,\omega)| \wedge |\tilde{X}^M(t,\omega)| \geq L\},$$

with $\inf \varnothing = \infty$. Observe that $\mathbb{P}(\tau_L \leq T) \to 0$ as $L \to \infty$ since $X^M, \tilde{X}^M$ are continuous processes.
Using Hölder's inequality, Doob's maximal inequality, the Itô isometry, and

condition (L), we obtain for $M \in \mathbb{N}$, all $t \in [0,T]$:

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{s\in[0,t]} |X^M_{s\wedge\tau_L} - \tilde{X}^M_{s\wedge\tau_L}|^2\Big] &\le 4T\mathbb{E}\left[\int_0^{t\wedge\tau_L} |b(r,X^M,\theta_r) - b(r,\tilde{X}^M,\theta_r)|^2 dr\right] \\
&\quad + 4\mathbb{E}\left[\int_0^{t\wedge\tau_L} |\sigma(r,X^M,\theta_r) - \sigma(r,\tilde{X}^M,\theta_r)|^2 dr \int_0^{t\wedge\tau_L} |u_M(r)|^2 dr\right] \\
&\quad + 16\mathbb{E}\left[\int_0^{t\wedge\tau_L} |\sigma(r,X^M,\theta_r) - \sigma(r,\tilde{X}^M,\theta_r)|^2 dr\right] \\
&\le 4T\mathbb{E}\left[\int_0^{t\wedge\tau_L} |b(r,X^M,\theta_r) - b(r,\tilde{X}^M,\theta_r)|^2 dr\right] \\
&\quad + (4M+16)\mathbb{E}\left[\int_0^{t\wedge\tau_L} |\sigma(r,X^M,\theta_r) - \sigma(r,\tilde{X}^M,\theta_r)|^2 dr\right] \\
&\le (4T+4M+16)\mathbb{E}\left[\int_0^{t\wedge\tau_L} \sup_{s\in[0,r]} |X^M_s - \tilde{X}^M_s|^2 dr\right] \\
&\le (4T+4M+16)\int_0^t \mathbb{E}\left[\sup_{s\in[0,r]} |X^M_{s\wedge\tau_L} - \tilde{X}^M_{s\wedge\tau_L}|^2\right] dr.
\end{aligned}
$$

An application of Gronwall's lemma yields that

$$
\mathbb{E}\Big[\sup_{s\in[0,T]} |X^M_{s\wedge\tau_L} - \tilde{X}^M_{s\wedge\tau_L}|^2\Big] = 0,
$$

hence, $\forall M \in \mathbb{N}$, $\mathbb{P}(X^M(t) = \tilde{X}^M(t), \text{ for all } t < \tau_L) = 1$ for all $L \in \mathbb{N}$. This implies the assertion since $\tau_L \nearrow \infty$ as $L \to \infty$ $\mathbb{P}$-almost surely.

Now, we can prove the existence part.
Fix the initial condition X_0 and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where it is defined a $\mathbb{R}^{d_1}$-valued Wiener process, $W = (W_t)_{t\in[0,T]}$. Choose the standard expansion $(\bar{\mathcal{G}}_{t+})_{t\in[0,T]}$ of the natural filtration generated by the Brownian Motion and the initial condition X_0. On this space, for each $M \in \mathbb{N}$, we will build a process X^M continuous and adapted which is a strong solution of the SDE (B.3) and we will show that $X^M \in M_2[0,T]$. We will get the process X^M, through an iterative process. Our procedure produce a solution X^M adapted to $(\bar{\mathcal{G}}_{t+})_{t\in[0,T]}$. Since $(\bar{\mathcal{G}}_{t+}) \subset (\mathcal{F}_t)$, X^M must be adapted to $(\mathcal{F}_t)_{t\in[0,T]}$, too. We have already shown uniqueness of solutions for (B.3). This implies that every solution Y defined on Ω, a priori adapted to $(\mathcal{F}_t)_{t\in[0,T]}$, is indistinguishable from X^M, hence, it is adapted to $(\bar{\mathcal{G}}_{t+})_{t\in[0,T]}$,too.

For $Y \in M_2[0,T]$, we define the process $J^M(Y) = (J_t^M(Y))_{t\in[0,T]}$ as:

$$J_t^M(Y) := X_0 + \int_0^t \sigma(s,Y,\theta_s)dW_s + \int_0^t b(s,Y,\theta_s)ds + \int_0^t \sigma(s,Y,\theta_s)u_M(s)ds.$$

Thanks to hypothesis (G), it follows immediately that drift and volatility terms are in $M_2[0,T]$, too. Exploiting Doob's maximal inequality, Itô isometry, Hölder' inequality and condition (G), we are able to write:

$$\begin{aligned}
\mathbb{E}\Big[\sup_{t\in[0,T]} |J_t^M(Y)|^2\Big] &\le 4\Bigg(\mathbb{E}[|X_0|^2] + 4\mathbb{E}\Big[\int_0^T |\sigma(s,Y,\theta_s)|^2 ds\Big] \\
&\quad + T\mathbb{E}\Big[\int_0^T |b(s,Y,\theta_s)|^2 ds\Big] + \mathbb{E}\Big[\int_0^T |\sigma(s,Y,\theta_s)|^2 ds \int_0^T |u_M(s)|^2 ds\Big]\Bigg) \\
&\le 4\Bigg(\mathbb{E}[|X_0|^2] + 4TK^2 + T\mathbb{E}\Big[\int_0^T |K(1+\sup_{r\in[0,s]}|Y_r|)|^2 ds\Big] + K^2T\mathbb{E}\Big[\int_0^T |u(s)|^2 ds\Big]\Bigg) \\
&\le 4\Bigg(\mathbb{E}[|X_0|^2] + 4TK^2 + 2K^2T^2\Bigg(\mathbb{E}\Big[\sup_{r\in[0,T]}|Y_r|^2\Big] + 1\Bigg) + K^2TQ\Bigg) < \infty.
\end{aligned}$$

This fact proves that $J^M(Y) \in M_2[0,T]$, for all $Y \in M_2[0,T]$, and so that $J^M(\cdot)$ is a functional from $M_2[0,T]$ in itself.
Let $Y, Y' \in M_2[0,T]$. Exploiting Doob's maximal inequality, Itô isometry, Hölder's inequality and condition (L), we can show:

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]} |J_t^M(Y) - J_t^M(Y')|^2\right] &\le 4T\mathbb{E}\left[\int_0^T |b(r,Y,\theta_r) - b(r,Y',\theta_r)|^2 dr\right] \\
&\quad + 4\mathbb{E}\left[\int_0^T |\sigma(r,Y,\theta_r) - \sigma(r,Y',\theta_r)|^2 dr \int_0^T |u_M(r)|^2 dr\right] \\
&\quad + 16\mathbb{E}\left[\int_0^T |\sigma(r,Y,\theta_r) - \sigma(r,Y',\theta_r)|^2 dr\right] \\
&\le 4T\mathbb{E}\left[\int_0^T |b(r,Y,\theta_r) - b(r,Y',\theta_r)|^2 dr\right] \\
&\quad + (4M+16)\mathbb{E}\left[\int_0^T |\sigma(r,Y,\theta_r) - \sigma(r,Y',\theta_r)|^2 dr\right] \\
&\le (4T+4M+16)\mathbb{E}\left[\int_0^T \sup_{s\in[0,r]} |Y_s - Y'_s|^2 dr\right] \\
&\le (4T+4M+16)\int_0^T \mathbb{E}\left[\sup_{s\in[0,r]} |Y_s - Y'_s|^2\right] dr.
\end{aligned}$$

Hence, we have the following:

$$\mathbb{E}\left[\sup_{t\in[0,T]}|J_t^M(Y)-J_t^M(Y')|^2\right] \le (4T+4M+16)\int_0^T \mathbb{E}\left[\sup_{s\in[0,r]}|Y_s-Y_s'|^2\right]dr$$
$$= C_M\int_0^T \mathbb{E}\left[\sup_{s\in[0,r]}|Y_s-Y_s'|^2\right]dr. \tag{B.4}$$

This fact naturally yields that

$$\mathbb{E}\left[\sup_{t\in[0,T]}|J_t^M(Y)-J_t^M(Y')|^2\right] \le C_M T\mathbb{E}\left[\sup_{s\in[0,T]}|Y_s-Y_s'|^2\right].$$

Hence, $J^M : M_2[0,T] \longrightarrow M_2[0,T]$ is a continuous functional.
Now, we define iteratively a sequence of processes $X^{(M,n)} = (X_t^{(M,n)})_{t\in[0,T]} \in M_2[0,T]$, as $X^{(M,1)} \equiv X_0$ and $X^{(M,n+1)} = J^M(X^{(M,n)}), n \in \mathbb{N}$.
To be more precise, for $t \in [0,T]$ and $n \in \mathbb{N}$, we define:

$$\begin{aligned} X_t^{(M,1)} &= X_0, \\ X_t^{(M,n+1)} &= J_t^M(X^{(M,n)}) = X_0 + \int_0^t \sigma(s, X^{(M,n)}, \theta_s)dW_s \\ &\quad + \int_0^t b(s, X^{(M,n)}, \theta_s)ds + \int_0^t \sigma(s, X^{(M,n)}, \theta_s)u_M(s)ds. \end{aligned}$$

Exploiting the relation (B.4), for all $n \ge 2$ and $t \in [0,T]$, we have:

$$\mathbb{E}\left[\sup_{t\in[0,T]}|X_t^{(M,n+1)}-X_t^{(M,n)}|^2\right] \le C_M\int_0^T \mathbb{E}\left[\sup_{t\in[0,r]}|X_t^{(M,n)}-X_t^{(M,n-1)}|^2\right]dr. \tag{B.5}$$

For $n = 1$, using Itô isometry, Cauchy-Schwarz inequality, Doob's maximal inequality and (G) hypothesis, we have:

$$\begin{aligned}
\mathbb{E}\Bigg[\sup_{t\in[0,T]}|X_t^{(M,2)}-X_t^{(M,1)}|^2\Bigg] &= \mathbb{E}\Bigg[\sup_{t\in[0,T]}|X_0+\int_0^t \sigma(s,X_0,\theta_s)dW_s \\
&+ \int_0^t b(s,X_0,\theta_s)ds + \int_0^t \sigma(s,X_0,\theta_s)u_M(s)ds - X_0|^2\Bigg] \\
&\le 12\mathbb{E}\Bigg[\int_0^T |\sigma(s,X_0,\theta_s)|^2 ds\Bigg] + 3T\mathbb{E}\Bigg[\int_0^T |b(s,X_0,\theta_s)|^2 ds\Bigg] \\
&+ 3\mathbb{E}\Bigg[\int_0^T |\sigma(s,X_0,\theta_s)|^2 ds \int_0^T |u_M(s)|^2 ds\Bigg] \\
&\le 12K^2T + 3T\mathbb{E}\Bigg[\int_0^T K^2(1+\sup_{r\in[0,s]}|X_0|)^2 ds\Bigg] + 3K^2TQ \\
&\le 12K^2T + 6T^2K^2(1+\mathbb{E}[|X_0|^2]) + 3K^2TQ.
\end{aligned}$$

Hence, we get:

$$\mathbb{E}\Bigg[\sup_{t\in[0,T]}|X_t^{(M,2)}-X_t^{(M,1)}|^2\Bigg] \le 12K^2T + 6T^2K^2(1+\mathbb{E}[|X_0|^2]) + 3K^2TQ = S. \tag{B.6}$$

Exploiting the relations (B.5) and (B.6), we can show by induction the following estimation for all $t \in [0,T]$ and $n \in \mathbb{N}$:

$$\mathbb{E}\Bigg[\sup_{t\in[0,T]}|X_t^{(M,n+1)}-X_t^{(M,n)}|^2\Bigg] \le SC_M^{n-1}\frac{T^{n-1}}{(n-1)!}. \tag{B.7}$$

This relation shows that the sequence of processes $(X^{(M,n)})_{n\in\mathbb{N}}$ is a Cauchy-sequence in $M_2[0,T]$. In fact, for $m > n$, we can write:

$$\begin{aligned}
\sqrt{\mathbb{E}\Bigg[\sup_{t\in[0,T]}|X_t^{(M,m)}-X_t^{(M,n)}|^2\Bigg]} &\le \sum_{k=n}^{m-1}\sqrt{\mathbb{E}\Bigg[\sup_{t\in[0,T]}|X_t^{(M,k+1)}-X_t^{(M,k)}|^2\Bigg]} \\
&\le \sum_{k=n}^{m-1}\sqrt{SC_M^{k-1}\frac{T^{k-1}}{(k-1)!}}
\end{aligned}$$

This sequence can be made, arbitrarily, small, choosing n big enough, since it is a convergent sequence. $M_2[0,T]$, with the metric introduced above, is a complete metric space.* Hence, there must exist a process $X^M \in M_2[0,T]$

*The proof of this fact follows immediately exploiting Lemma 4.6 in [9].

such that $X^{(M,n)} \to_{n\to\infty} X^M$ in $M_2[0,T]$.
Finally, let's show that X^M is a solution for equation (B.3).
By construction, we have $X^{(M,n+1)} = J^M(X^{(M,n)})$, for all $n \in \mathbb{N}$. Since $X^{(M,n)} \to X^M$ in $M_2[0,T]$, taking the limit for $n \to \infty$ and exploiting the fact that $J^M : M_2[0,T] \to M_2[0,T]$ is a continuous operator, we get the relation $X^M = J^M(X^M)$, that is:

$$X_t^M = X_0 + \int_0^t b(s, X^M, \theta_s)ds + \int_0^t \sigma(s, X^M, \theta_s)u_M(s)ds + \int_0^t \sigma(s, X^M, \theta_s)dW_s, \tag{B.8}$$

which is the integral form of equation (B.3).
By construction, X^M has been determined as an element of $M_2[0,T]$, namely as an equivalence class of processes. We need to show that it is possible to choose a real process X^M, so an element in the equivalence class, that is continuous. The expression (B.8) shows that X^M is the sum of a stochastic integral and of two ordinary integral, hence, there must be a continuous version of it: this version of X^M is a solution for equation (B.3). This must be a strong solution of (B.3), since at the beginning we decided to work with the completion of the natural filtration of the Brownian motion, $(\bar{\mathcal{G}_{t+}})_{t\in[0,T]}$.
We have shown that $\forall M \in \mathbb{N}$ there exists a unique strong solution X^M of equation (B.3).
We can define the process $X = (X_t)_{t\in[0,T]}$ as:

$$X(t,\omega) = \begin{cases} X^M(t,\omega) & \text{if } t \in (\tau_{M-1}(\omega), \tau_M(\omega)], \qquad \forall M \in \mathbb{N}_0 \\ X_0(\omega) & \text{if } t = 0 \end{cases}$$

The process X is well defined since, for $t \in [0, \tau_M]$, we have $X^M(t) = X^{M+1}(t)$, $\mathbb{P} - a.s.$. In fact, for $t \in [0, \tau_M]$, $u_M(t) = u_{M+1}(t) = u(t)$ and so the two processes solve the same SDE that has a unique strong solution and therefore they must be indistinguishable.
Now, we need to show that X is a solution of equation (B.1).
Exploiting the fact that for $s \in [0, \tau_M]$ $\quad X_s^M = X_s$, $\forall M \in \mathbb{N}, \forall t \in [0,T]$, we have

$$\begin{aligned} X(t\wedge\tau_M) &= X^M(t\wedge\tau_M) \\ &= X_0 + \int_0^{t\wedge\tau_M} b(s, X^M, \theta_s)ds + \int_0^{t\wedge\tau_M} \sigma(s, X^M, \theta_s)u(s)\mathbb{I}_{[0,\tau_M]}(s)ds \\ &\quad + \int_0^{t\wedge\tau_M} \sigma(s, X^M, \theta_s)dW_s \\ &= X_0 + \int_0^{t\wedge\tau_M} b(s, X, \theta_s)ds + \int_0^{t\wedge\tau_M} \sigma(s, X, \theta_s)u(s)ds \\ &\quad + \int_0^{t\wedge\tau_M} \sigma(s, X, \theta_s)dW_s. \end{aligned}$$

Since $\tau_M \to \infty$, as $M \to \infty$, in the limit we get:

$$X(t) = X_0 + \int_0^t b(s, X, \theta_s)ds + \int_0^t \sigma(s, X, \theta_s)u(s)ds \\ + \int_0^t \sigma(s, X, \theta_s)dW_s,$$

which is an integral version of equation (B.1).
By the uniqueness of solution for equation (B.1), we can conclude that X is the strong solution of (B.1) we were looking for.

Finally we can show the estimation (B.2).
Suppose that $(X_t)_{t\in[0,T]}$ is a solution for the equation (B.1), with $\theta \in \mathcal{C}([0,T], \mathcal{P}_1(\mathbb{R}^d))$ fixed. For $M \in \mathbb{N}$, define a $(\mathcal{F}_t)-$stopping time τ_M, by:

$$\tau_M(\omega) := \inf\{t \in [0,T] : |X(t,\omega)| \geq M\},$$

with $\inf \varnothing = +\infty$.
Let $t \in [0,T]$, exploiting in order Hölder's inequality, Itô's isometry and Doob's maximal inequality, we have:

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\in[0,t]} |X_{s\wedge\tau_M}|^2\Big] &\leq 4\mathbb{E}\big[|X_0|^2\big] + 4\mathbb{E}\left[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta_r)|^2 dr \int_0^{t\wedge\tau_M} |u(r)|^2 dr\right] \\ &+ 4T\mathbb{E}\left[\int_0^{t\wedge\tau_M} |b(r,X,\theta_r)|^2 dr\right] + 16\mathbb{E}\left[|\int_0^{t\wedge\tau_M} \sigma(r,X,\theta_r)dW_r|^2\right] \\ &\leq 4\mathbb{E}\big[|X_0|^2\big] + 4\mathbb{E}\left[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta_r)|^2 dr \int_0^t |u(r)|^2 dr\right] \\ &+ 4T\mathbb{E}\left[\int_0^{t\wedge\tau_M} |b(r,X,\theta_r)|^2 dr\right] + 16\mathbb{E}\left[\int_0^{t\wedge\tau_M} |\sigma(r,X,\theta_r)|^2 dr\right] = \diamondsuit\end{aligned}$$

Exploiting (G) hypotheses on b and σ:

$$\diamondsuit \leq 4\mathbb{E}\big[|X_0|^2\big] + 4T\mathbb{E}\Bigg[\int_0^{t\wedge\tau_M} K^2(1+\sup_{s\in[0,r]}|X_s|)^2 dr\Bigg] + 4K^2T\mathbb{E}\Bigg[\int_0^t |u(r)|^2 dr\Bigg]$$
$$+ 16K^2T$$
$$\leq 4\mathbb{E}\big[|X_0|^2\big] + 4T\mathbb{E}\Bigg[\int_0^{t\wedge\tau_M} 2K^2(1+\sup_{s\in[0,r]}|X_s|^2) dr\Bigg] + 4K^2T\mathbb{E}\Bigg[\int_0^t |u(r)|^2 dr\Bigg]$$
$$+ 16K^2T$$
$$\leq 4\mathbb{E}\big[|X_0|^2\big] + 8TK^2\int_0^t \mathbb{E}\Bigg[\sup_{s\in[0,r]}|X_{s\wedge\tau_M}|^2\Bigg] dr + 8T^2K^2 + 4K^2T\mathbb{E}\Bigg[\int_0^t |u(r)|^2 dr\Bigg]$$
$$+ 16K^2T$$
$$= R + 8TK^2\int_0^t \mathbb{E}\Bigg[\sup_{s\in[0,r]}|X_{s\wedge\tau_M}|^2\Bigg] dr.$$

Now, applying Gronwall Lemma, we have that for each $M \in \mathbb{N}$, for each $t \in [0,T]$:

$$\mathbb{E}\Big[\sup_{s\in[0,t]}|X_{s\wedge\tau_M}|^2\Big] \leq Re^{8TK^2t}.$$

In particular, for $t = T$, we have:

$$\mathbb{E}\Big[\sup_{s\in[0,T]}|X_{s\wedge\tau_M}|^2\Big] \leq Re^{8(TK)^2} = R' < \infty.$$

Finally, exploiting monotone convergence theorem, we get:

$$\mathbb{E}\Big[\sup_{s\in[0,T]}|X_s|^2\Big] = \lim_{M\to\infty}\mathbb{E}\Big[\sup_{s\in[0,T]}|X_{s\wedge\tau_M}|^2\Big] \leq Re^{8(TK)^2} = R' < \infty.$$

□

Remark B.3. *Now, consider the McKean-Vlasov Equation:*

$$dX_t = b(t,X,Law(X(t)))dt + \sigma(t,X,Law(X(t)))u(t)dt + \sigma(t,X,Law(X(t)))dW_t,$$
$$X(0) = X_0 \tag{B.9}$$

with the same hypotheses on coefficients and control as in the Theorem (B.1). If $\tilde{X}$ is a solution of Equation (B.9), i.e. a solution of Equation (B.1) with $\theta_t = Law(X_t)$, we are able to show the finiteness of $\mathbb{E}\Big[\sup_{t\in[0,T]}|\tilde{X}_t|^2\Big]$, with the same arguments used in the proof above.

Bibliography

[1] Aliprantis, Charalambos D.; Border, Kim C.(2006-3rd edition): *Infinite Dimensional Analysis*, Springer.

[2] Billingsley, Patrick (1999-2nd edition): *Convergence of Probability Measures*,Wiley series in Probability and Statistics, New York.

[3] Budhiraja, Amarjit; Dupuis, Paul; Fischer, Markus (2012): *Large Deviation properties of weakly interacting processes via weak convergence methods*, The Annals of Probability, vol. 40, No.I, 74-102.

[4] Dawson, Donald A; Gartner, Jurgen (1987): *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, Stochastics: formerly Stochastics and Stochastics Reports, 20:4, 247-308.

[5] Dobrushin, R. L.; (1970): *Prescribing a system of random variables by conditional distributions*,Th. Probab. and its Applic., 3, 469 (1970).

[6] Dudley, R. M.; (1989): *Real Analysis and Probability*,Cambridge Studies in Advanced Probability, Cambridge.

[7] Fischer, Markus (2014): *On the form of the large deviation rate function for the empirical measures of weakly interacting systems*, Bernoulli, 20(4), 1765-1801.

[8] Kallenberg, Olav (1996): *On the existence of Universal Functional Solutions to Classical SDE's*,The Annals of Probability, vol. 24, No.1, 196-205.

[9] Kallenberg, Olav (1997): *Foundations of Modern Probability*,Springer.

[10] Karatzas, I. and Shreve, S. E. (1991): *Brownian Motion and Stochastic Calculus*, 2nd ed. Graduate Texts in Mathematics 113. Springer, New York.

[11] McKean, H. P. (1966): *A class of Markov processes associated with nonlinear parabolic equations*, Proceedings of the National Academy of Sciences, 56, 1907–1911.

[12] Oelschlager, Karl (1984): *A Martingale approach to the Law of Large Nmbers for Weakly Interacting Stochastic Processes*, The Annals of Probability, vol. 12, No.2, 458-479.

[13] Scheutzow, Michael (1987):*Uniqueness and Non-uniqueness of solutions of Vlasov-McKean Equations*,J. Australian Mathematical Society (Series A), vol. 43, 246-256 .

[14] Sznitman, Alain-Sol (1991): *Topics in propagation of chaos*, Hennequin PL. (eds) Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989. Lecture Notes in Mathematics, vol 1464. Springer, Berlin, Heidelberg

[15] Villani, Cedric; (2003):*Topics in Optimal Transportation*,American Mathematical Society, Graduate Studies in Mathematics, vol 58 .